(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
29 August 2002 (29.08.2002)

PCT

(10) International Publication Number
WO 02/066501 A2

(51) International Patent Classification⁷: C07K 14/205, C12N 15/31, G01N 33/68, C12N 15/63, A61K 39/106, 35/74

(21) International Application Number: PCT/EP01/15428

(22) International Filing Date:
28 December 2001 (28.12.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/259,302 2 January 2001 (02.01.2001) US

(71) Applicants *(for all designated States except US)*: **HYBRIGENICS** [FR/FR]; 3/5 Impasse Reille, F-75014 Paris (FR). **INSTITUT PASTEUR** [FR/FR]; 25-28, rue du Dr Roux, F-75724 Paris Cedex 15 (FR).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **LEGRAIN, Pierre** [FR/FR]; 5, rue Mizon, F-75015 Paris (FR). **RAIN, Jean-Christophe** [FR/FR]; 32, Jardin Boeldieu, F-92800 Puteaux (FR). **COLLAND, Frédéric** [FR/FR]; 16, rue du Manoir, F-95380 Puiseux en France (FR). **DE REUSE, Hilde** [BE/FR]; 49, rue Rouelle, F-75015 Paris (FR). **LABIGNE, Agnès** [FR/FR]; 29, rue du Bois Michel Pierre, Appt 106, F-91440 Bures sur Yvette (FR).

(74) Agent: **ERNEST GUTMANN - YVES PLASSERAUD S.A.**; 3, rue Chauveau-Lagarde, F-75008 Paris (FR).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: PROTEIN-PROTEIN INTERACTIONS IN *HELICOBACTER PYLORI*

(57) Abstract: The present invention relates to protein-protein interactions of *Helicobacter pylori*. More specifically, the present invention relates to complexes of polypeptides or polynucleotides encoding the polypeptides, fragments of the polypeptides, antibodies to the complexes, Selected Interacting Domains (SID®) which are identified due to the protein-protein interactions, methods for screening drugs for agents which modulate the interaction of proteins and pharmaceutical compositions that are capable of modulating the protein-protein interactions.

WO 02/066501 A2

PROTEIN-PROTEIN INTERACTIONS IN *Helicobacter pylori*

FIELD OF THE INVENTION

The present invention relates to proteins that interact with *Helicobacter* pylori. More specifically, the present invention relates to complexes of polypeptides or polynucleotides encoding the polypeptides, fragments of the polypeptides, antibodies to the complexes, Selected Interacting Domains (SID®) which are identified due to the protein-protein interactions, methods for screening drugs for agents which modulate the interaction of proteins and pharmaceutical compositions that are capable of modulating the protein-protein interactions.

In another embodiment the present invention provides a protein-protein interaction map called a PIM® which is available in a report relating to the protein-protein interactions of *Helicobacter pylori*.

BACKGROUND AND PRIOR ART

Most biological processes involve specific protein-protein interactions. Protein-protein interactions enable two or more proteins to associate. A large number of non-covalent bonds form between the proteins when two protein surfaces are precisely matched. These bonds account for the specificity of recognition. Thus, protein-protein interactions are involved, for example, in the assembly of enzyme subunits, in antibody-antigen recognition, in the formation of biochemical complexes, in the correct folding of proteins, in the metabolism of proteins, in the transport of proteins, in the localization of proteins, in protein turnover, in first translation modifications, in the core structures of viruses and in signal transduction.

General methodologies to identify interacting proteins or to study these interactions have been developed. Among these methods are the two-hybrid system originally developed by Fields and co-workers and described, for

example, in U.S. Patent Nos. 5,283,173, 5,468,614 and 5,667,973, which are hereby incorporated by reference.

The earliest and simplest two-hybrid system, which acted as basis for development of other versions, is an *in vivo* assay between two specifically constructed proteins. The first protein, known in the art as the "bait protein" is a chimeric protein which binds to a site on DNA upstream of a reporter gene by means of a DNA-binding domain or BD. Commonly, the binding domain is the DNA-binding domain from either Gal4 or native *E. coli* LexA and the sites placed upstream of the reporter are Gal4 binding sites or LexA operators, respectively.

The second protein is also a chimeric protein known as the "prey" in the art. This second chimeric protein carries an activation domain or AD. This activation domain is typically derived from Gal4, from VP 16 or from B42.

Besides the two hybrid systems, other improved systems have been developed to detect protein-protein interactions. For example, a two-hybrid plus one system was developed that allows the use of two proteins as bait to screen available cDNA libraries to detect a third partner. This method permits the detection between proteins that are part of a larger protein complex such as the RNA polymerase II holoenzyme and the TFIIH or TFIID complexes. Therefore, this method, in general, permits the detection of ternary complex formation as well as inhibitors preventing the interaction between the two previously defined fused proteins.

Another advantage of the two-hybrid plus one system is that it allows or prevents the formation of the transcriptional activator since the third partner can be expressed from a conditional promoter such as the methionine-repressed Met25 promoter which is positively regulated in medium lacking methionine. The presence of the methionine-regulated promoter provides an excellent control to evaluate the activation or inhibition properties of the third partner due to its "on" and "off" switch for the formation of the transcriptional activator. The three-hybrid method is described, for example in Tirode et al., *The Journal of Biological*

Chemistry, **272,** No. 37 pp. 22995-22999 (1997) incorporated herein by reference.

Besides the two and two-hybrid plus one systems, yet another variant is that described in Vidal et al, Proc. *Natl. Sci.* **93** pgs. 10315-10320 called the reverse two-and one-hybrid systems where a collection of molecules can be screened that inhibit a specific protein-protein or protein/DNA interaction, respectively.

A summary of the available methodologies for detecting protein-protein interactions is described in Vidal and Legrain, *Nucleic Acids Research* Vol. 27, No. 4 pgs.919-929 (1999) and Legrain and Selig, FEBS Letters 480 pgs. 32-36 (2000) which references are incorporated herein by reference.

However, the above conventionally used approaches and especially the commonly used two-hybrid methods have their drawbacks. For example, it is known in the art that, more often than not, false positives and false negatives exist in the screening method. In fact, a doctrine has been developed in this field for interpreting the results and in common practice an additional technique such as co-immunoprecipitation or gradient sedimentation of the putative interactors from the appropriate cell or tissue type are generally performed. The methods used for interpreting the results are described by Brent and Finley, Jr. in *Ann. Rev. Genet.,* **31** pgs. 663-704 (1997). Thus, the data interpretation is very questionable using the conventional systems.

One method to overcome the difficulties encountered with the methods in the prior art is described in WO 99/42612, incorporated herein by reference. This method is similar to the two-hybrid system described in the prior art in that it also uses bait and prey polypeptides. However, the difference with this method is that a step of mating at least one first haploid recombinant yeast cell containing the prey polypeptide to be assayed with a second haploid recombinant yeast cell containing the bait polynucleotide is performed. Of course the person skilled in the art would appreciate that either the first recombinant yeast cell or the second

recombinant yeast cell also contains at least one detectable reporter gene that is activated by a polypeptide including a transcriptional activation domain.

The method described in W099/42612 permits the screening of more prey polynucleotides with a given bait polynucleotide in a single step than in the prior art systems due to the cell to cell mating strategy between haploid yeast cells. Furthermore, this method is more thorough and reproducible, as well as sensitive. Thus, the presence of false negatives and/or false positives is extremely minimal as compared to the conventional prior art methods.

One of the prokaryotic microorganisms studied by the inventors is *Helicobacter pylori*. *Helicobacter pylori (H. pylori)* is a microaerophilic, Gram negative, slow growing, spiral shaped and flagellated organism. *H. pylori* has been first isolated in 1983 from a gastric biopsy specimen of patient with chronic gastritis (Marshall et al., 1984, Lancet, 1:1311-1314, *Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration).*

Helicobacter pylori has become identified as a primary cause of chronic gastroduodenal disorders, such as gastritis, dyspepsia, and peptic ulcers, in humans. Studies have shown (Labigne et al.) that *H. pylori* can be successfully eradicated by a treatment combining two antibiotics with a proton pump inhibitor. However, few antibiotics are active against *H. pylori* and antibiotic-resistant strains have begun to appear.

The *H. pylori* strain n° 26695 genome has been studied by Tomb et al. (Tomb et al., 1997, Nature, vol. 388, 539-547, *The complete genome sequence of the gastric pathogen Helicobacter pylori).* This strain's genome consists of a circular chromosome with a size of 1,667,867 bp, average G + C content of 39%, and 1590 predicted coding sequences (open reading frames or "ORF").

The availability of the entire genome sequence of two clinical strains, 26695 and J99 (Alm et al, 1999) (Tomb et al, 1997) has encouraged more global approaches to the functional analysis of the whole set of genes. In *H. pylori*, there is still as much as 42% of the *H. pylori* encoded proteins for which there is a need to assign biological function. Attempts to classify genes within functional

categories such as genes essential for viability, or conditionally essential in a given environment have been proposed by Akerley and collaborators (Akerley et al, 1998). This concept was first applied to *H. pylori* by Chalker et al. (Chalker et al, 2001) who analyzed for essentiality a set of genes selected by bioinformatic genome prioritization.

The bacterial factors necessary for colonization of the gastric environment, and for virulence of this pathogen, are poorly understood. Examples of known virulence factors are:

- Enzymes involved in neutralizing the acid gastric pH : the multisubunit urease is a characteristic enzyme that is crucial for survival in acidic pH and for successful colonization of the gastric environment, a site that few other microbes can colonize (Labigne et al., WO 93/07273, *Helicobacter pylori genes necessary for the regulation and maturation of urease, and use thereof*). Genes encoding ureases have been located on a 34 kb chromosome fragment and comprise ureA, ureB, ureC, ureD, ureE, ureF, ureG, ureH and ureI.

- Bacterial flagellar proteins responsible for motility across the mucous layer (Hazell et al., 1986, J. Inf. Dis., 153, 658-663 *Campylobacter pyloridis and gastritis: association with intracellular spaces and adaptation to an environment of mucus as important factors in colonization of the gastric epithelium;* Leying et al., 1992, Mol. Microbiol., 6, 2863-2874 *Cloning and genetic characterization of Helicobacter pylori flagellin gene):* flagellar filaments biosynthesis comprises A and B flagellins and the filament cap. These two biosyntheses are regulated by flbA gene (Suerbaum et al., French patent application, 1995, n° 2,736,360, *Cloning and characterization of flbA gene of Helicobacter* pylori, *aflagellated strains production).*

- Two other essential toxins for virulence are VacA and CagA.

VacA is a *H. pylori* toxin that induces the formation of large acidic vacuoles in host epithelial cells. These large vacuoles originate from massive swelling of

membranous compartments of late stages of the endocytic pathway (de Bernard et al., 1997, Microbiology, 26(4), 665-674, *Helicobacter* pylori *toxin VacA induces vacuole formation by acting in the cell cytosol)* Proof for receptor-mediated interaction with VacA has been made by Pagliaccia et al.; m2 allele of vacA gene has always been described as inactive in the *in vitro* HeLa cell assay, however, the m2 allele is associated with peptic ulcer and is prevalent in populations in which peptic ulcer and gastric cancer have high incidence (Pagliaccia et al., Proc. Natl. Acad. Sci. U.S.A, 1998, 95(17), 10212-10217, *The m2 form of the Helicobacter* pylori *cytotoxin has cell type-specific vacuolating activity).*

CagA is one of the proteins encoded by the "cag pathogenicity island" (Spohn et al. 1997, Molecular Microbiology, 26(2), 361-372,. *Transcriptional analysis of the divergent cagAB genes encoded by the pathogenicity island of Helicobacter pylori)* found in *H. pylori* strains isolated from most patients with peptic ulcer disease and adenocarcinoma. CagA is produced by 50-60% of *H. pylori* strains; it is a high molecular weight (120-140 kDa) superficial protein and an immunodominant antigen with unknown function. *H. pylori* strains that produce CagA protein have two genes cagB and cagC (36 and 101 kDa proteins, respectively). These genes are highly associated with duodenal ulcers (Blaser et al. 1996, WO 96/12825, *cagB and cagC genes of Helicobacter pylori and related methods an d compositions).*

Other virulence factors are : several gastric tissue-specific adhesins (Boren et al., 1993, Science, **262**, 1892-1895).

Therapeutic agents are currently available that eradicate *H. pylori* infections *in vitro*. However, methods employing antibiotic agents result in the emergence of bacterial strains which are resistant to these agents.

Thus, it is an object of the present invention to identify protein-protein interactions for *Helicobacter pylori.*

It is another object of the present invention to identify protein-protein interactions of *Helicobacter pylori* for the development of more effective and better targeted therapeutic applications.

It is yet another object of the present invention to identify complexes of polypeptides or polynucleotides encoding the polypeptides and fragments of the polypeptides of *Helicobacter pylori*.

It is yet another object of the present invention to identify antibodies to these complexes of polypeptides or polynucleotides encoding the polypeptides and fragments of the polypeptides of *Helicobacter pylori* including polyclonal, as well as monoclonal antibodies that are used for detection.

It is still another object of the present invention to identify selected interacting domains of the polypeptides, called SID® polypeptides.

It is still another object of the present invention to identify selected interacting domains of the polynucleotides, called SID® polynucleotides.

It is another object of the present invention to generate protein-protein interactions maps called PIM®s.

It is yet another object of the present invention to classify genes of *H. pylori* into functional categories at the genomic scale such as genes essential or nonessential for viability.

It is yet another object of the present invention to identify Putative Essential Genes (PEGS) from *H. pylori*, as well as the true essential genes.

It is yet another object of the present invention to establish a large scale protein-protein interaction map of *H. pylori* using the two-hybrid system as a way to elucidate the function of yet uncharacterized proteins.

It is yet another object of the present invention to identify a superbinder phenotype in *H. pylori* with the two-hybrid system which completely inhibits specific protein-protein interactions.

It is yet another object of the present invention to identify oligopeptides. Their overlapping or combining derivatives that inhibit *H. pylori* growth.

In yet another aspect, the present invention relates to the identification of ORFs (open reading frames) having enzymatic activity, which provides a direct way to screen lead compounds that abolish enzymatic activity through the disruption of the oligomeric interaction.

It is yet another object of the present invention to provide a method for screening drugs for agents which modulate the interaction of proteins and pharmaceutical compositions that are capable of modulating the protein-protein interactions of *Helicobacter pylori*.

It is another object to administer the nucleic acids of the present invention via gene therapy.

It is yet another object of the present invention to provide protein chips or protein microarrays.

It is yet another object of he present invention to provide a report in, for example paper, electronic and/or digital forms, concerning the protein-protein interactions, the modulating compounds and the like, as well as a PIM®.

These and other objects are achieved by the present invention as evidenced by the summary of the invention, description of the preferred embodiments and the claims.

SUMMARY OF THE PRESENT INVENTION

Thus the present invention relates to a protein complex of polypeptides as described in Table 1.

Furthermore, the present invention provides SID® polynucleotides and SID® polypeptides as defined in Figure 2, as well as a PIM® for *Helicobacter pylori*.

The present invention also provides antibodies to the protein-protein complexes for *Helicobacter pylori*.

In another embodiment the present invention provides a method for screening drugs for agents that modulate the protein-protein interactions and pharmaceutical compositions that are capable of modulating protein-protein interactions.

In another embodiment the present invention provides protein chips or protein microarrays.

In another embodiment the present invention identifies a superbinder phenotype in H. pylori with the two-hybrid system which completely inhibits protein-protein interactions.

In another aspect the present invention provides oligopeptides, their overlapping or combining derivatives thereof that inhibit *H. pylori* growth.

In yet another embodiment, the present invention identifies ORFs having enzymatic activity which provides a direct way to screen lead compounds.

In yet another embodiment the present invention provides a report in, for example, paper, electronic and/or digital forms.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a schematic representation of the pB1 plasmid.

Fig. 2 is a schematic representation of the pB5 plasmid.

Fig. 3 is a schematic representation of the pB6 plasmid.

Fig. 4 is a schematic representation of the pB13 plasmid.

Fig. 5 is a schematic representation of the pB14 plasmid.

Fig. 6 is a schematic representation of the pB20 plasmid.

Fig. 7 is a schematic representation of the pP1 plasmid.

Fig. 8 is a schematic representation of the pP2 plasmid.

Fig. 9 is a schematic representation of the pP3 plasmid.

Fig. 10 is a schematic representation of the pP6 plasmid.

Fig. 11 is a schematic representation of the pP7 plasmid.

Fig. 12 is a schematic representation of vectors expressing the T25 fragment.

Fig. 13 is a schematic representation of vectors expressing the T18 fragment.

Fig. 14 is a schematic representation of various vectors of pCmAHL1, pT25 and pT18.

Fig. 15 is a schematic representation identifying the SID®'s of *Helicobacter* pylori. In this figure the "Full-length prey protein" is the Open Reading Frame (ORF) or coding sequence (CDS) where the identified prey polypeptides are included. The Selected Interaction Domain (SID®) is determined by the commonly shared polypeptide domain of every selected prey fragment.

Fig. 16 is a protein map (PIM®).

Fig. 17 is a gel illustrating the results obtained for the disruption of the ORFs *hp0099* to *hp0198*. This figure exemplifies first that multiple insertions of the transposon took place and second that for the majority trransposon insertion occurred at a distance ranging between 100 to 600 bp from the 5'-end of the ORF, a distance compatible with the promotion of gene replacement by allelic recombination.

Fig. 18 is a schematic diagram of the procedure for classification of the genes as described in the present invention.

Fig. 19 are the results of three-hybrid experiments. Growth phenotypes of diploid strains containing various plasmids were analyzed by incubating cells at various dilutions (from 1 to 10^{-4}). Yeast growth was performed

over 2 days at 30°C on DO-3+Met or DO-3-Met medium. Lane 1 are cells containing [p3H1-HP1230]+pP6-HP1529; Lane 2 [p3H1-HP1230-SID1529WT]+pP6-HP1529; lane 3[pH1-HP1230-SID1529(N38D-V53L)+pP6-HP1529; lane 4 [p3H1-HP1230-SID1529(V53L)]+pP6-HP1529; and lane 5 [p3H1-HP1230]+pP6-HP0875.

Fig. 20 is the pP7-centro vector.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

As used herein the terms "polynucleotides", "nucleic acids" and "oligonucleotides" are used interchangeably and include, but are not limited to RNA, DNA, RNA/DNA sequences of more than one nucleotide in either single chain or duplex form. The polynucleotide sequences of the present invention may be prepared from any known method including, but not limited to, any synthetic method, any recombinant method, any ex vivo generation method and the like, as well as combinations thereof.

The term "polypeptide" means herein a polymer of amino acids having no specific length. Thus, peptides, oligopeptides and proteins are included in the definition of "polypeptide" and these terms are used interchangeably throughout the specification, as well as in the claims. The term "polypeptide" does not exclude posttranslational modifications such as polypeptides having covalent attachment of glycosyl groups, acetyl groups, phosphate groups, lipid groups and the like. Also encompassed by this definition of "polypeptide" are homologs thereof.

By the term "homologs" is meant structurally similar genes contained within a given species, orthologs are functionally equivalent genes from a given species or strain, as determined for example, in a standard complementation assay. Thus, a polypeptide of interest can be used not only as a model for identifying similar genes in given strains, but also to identify homologs and orthologs of the polypeptide of interest in other species. The orthologs, for

example, can also be identified in a conventional complementation assay. In addition or alternatively, such orthologs can be expected to exist in bacteria (or other kind of cells) in the same branch of the phylogenic tree, as set forth, for example, at ftp://ft.cme.msu.edu/pub/rdp/SSU-rRNA/SSU/Prok.phylo.

As used herein the term "prey polynucleotide" means a chimeric polynucleotide encoding a polypeptide comprising (i) a specific domain; and (ii) a polypeptide that is to be tested for interaction with a bait polypeptide. The specific domain is preferably a transcriptional activating domain.

As used herein, a "bait polynucleotide" is a chimeric polynucleotide encoding a chimeric polypeptide comprising (i) a complementary domain; and (ii) a polypeptide that is to be tested for interaction with at least one prey polypeptide. The complementary domain is preferably a DNA-binding domain that recognizes a binding site that is further detected and is contained in the host organism.

As used herein "complementary domain" is meant a functional constitution of the activity when bait and prey are interacting; for example, enzymatic activity.

As used herein "specific domain" is meant a functional interacting activation domain that may work through different mechanisms by interacting directly or indirectly through intermediary proteins with RNA polymerase II or III-associated proteins in the vicinity of the transcription start site.

As used herein the term "complementary" means that, for example, each base of a first polynucleotide is paired with the complementary base of a second polynucleotide whose orientation is reversed. The complementary bases are A and T (or A and U) or C and G.

The term "sequence identity" refers to the identity between two peptides or between two nucleic acids. Identity between sequences can be determined by comparing a position in each of the sequences which may be aligned for the purposes of comparison. When a position in the compared sequences is occupied by the same base or amino acid, then the sequences are identical at that position. A degree of sequence identity between nucleic acid sequences is a

function of the number of identical nucleotides at positions shared by these sequences. A degree of identity between amino acid sequences is a function of the number of identical amino acid sequences that are shared between these sequences. Since two polypeptides may each (i) comprise a sequence (i.e., a portion of a complete polynucleotide sequence) that is similar between two polynucleotides, and (ii) may further comprise a sequence that is divergent between two polynucleotides, sequence identity comparisons between two or more polynucleotides over a "comparison window" refers to the conceptual segment of at least 20 contiguous nucleotide positions wherein a polynucleotide sequence may be compared to a reference nucleotide sequence of at least 20 contiguous nucleotides and wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) of 20 percent or less compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences.

To determine the percent identity of two amino acids sequences or two nucleic acid sequences, the sequences are aligned for optimal comparison. For example, gaps can be introduced in the sequence of a first amino acid sequence or a first nucleic acid sequence for optimal alignment with the second amino acid sequence or second nucleic acid sequence. The amino acrid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, the molecules are identical at that position.

The percent identity between the two sequences is a function of the number of identical positions shared by the sequences. Hence % identity = number of identical positions / total number of overlapping positions X 100.

In this comparison the sequences can be the same length or may be different in length. Optimal alignment of sequences for determining a comparison window may be conducted by the local homology algorithm of Smith and Waterman (J. *Theor. Biol, 91 (2)* pgs. *370-380 (1981),* by the homology alignment algorithm of Needleman and Wunsch, *J. Miol. Biol, 48(3)* pgs. *443-453*

(1972), by the search for similarity via the method of Pearson and Lipman, PNAS, USA, *85(5)* pgs. *2444-2448 (1988)*, by computerized implementations of these algorithms (GAP, BESTFIT, FASTA and TFASTA in the Wisconsin Genetics Software Package Release *7.0*, Genetic Computer Group, *575*, Science Drive, Madison, Wisconsin) or by inspection.

The best alignment (i.e., resulting in the highest percentage of identity over the comparison window) generated by the various methods is selected.

The term "sequence identity" means that two polynucleotide sequences are identical (i.e., on a nucleotide by nucleotide basis) over the window of comparison. The term "percentage of sequence identity" is calculated by comparing two optimally aligned sequences over the window of comparison, determining the number of positions at which the identical nucleic acid base (e.g., A, T, C, G, U, or I) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison (i.e., the window size) and multiplying the result by 100 to yield the percentage of sequence identity. The same process can be applied to polypeptide sequences.

The percentage of sequence identity of a nucleic acid sequence or an amino acid sequence can also be calculated using BLAST software (Version 2.06 of September 1998) with the default or user defined parameter.

The term "sequence similarity" means that amino acids can be modified while retaining the same function. It is known that amino acids are classified according to the nature of their side groups and some amino acids such as the basic amino acids can be interchanged for one another while their basic function is maintained.

The term "isolated" as used herein means that a biological material such as a nucleic acid or protein has been removed from its original environment in which it is naturally present. For example, a polynucleotide present in a plant, mammal or animal is present in its natural state and is not considered to be isolated. The same polynucleotide separated from the adjacent nucleic acid

sequences in which it is naturally inserted in the genome of the plant or animal is considered as being "isolated."

The term "isolated" is not meant to exclude artificial or synthetic mixtures with other compounds, or the presence of impurities which do not interfere with the biological activity and which may be present, for example, due to incomplete purification, addition of stabilizers or mixtures with pharmaceutically acceptable excipients and the like.

"Isolated polypeptide" or "isolated protein" as used herein means a polypeptide or protein which is substantially free of those compounds that are normally associated with the polypeptide or protein in a naturally state such as other proteins or polypeptides, nucleic acids, carbohydrates, lipids and the like.

The term "purified" as used herein means at least one order of magnitude of purification is achieved, preferably two or three orders of magnitude, most preferably four or five orders of magnitude of purification of the starting material or of the natural material. Thus, the term "purified" as utilized herein does not mean that the material is 100% purified and thus excludes any other material.

The term "variants" when referring to, for example, polynucleotides encoding a polypeptide variant of a given reference polypeptide are polynucleotides that differ from the reference polypeptide but generally maintain their functional characteristics of the reference polypeptide. A variant of a polynucleotide may be a naturally occurring allelic variant or it may be a variant that is known naturally not to occur. Such non-naturally occurring variants of the reference polynucleotide can be made by, for example, mutagenesis techniques, including those mutagenesis techniques that are applied to polynucleotides, cells or organisms.

Generally, differences are limited so that the nucleotide sequences of the reference and variant are closely similar overall and, in many regions identical.

Variants of polynucleotides according to the present invention include, but are not limited to, nucleotide sequences which are at least 95% identical after

alignment to the reference polynucleotide encoding the reference polypeptide. These variants can also have 96%, 97%, 98% and 99.999% sequence identity to the reference polynucleotide.

Nucleotide changes present in a variant polynucleotide may be silent, which means that these changes do not alter the amino acid sequences encoded by the reference polynucleotide.

Substitutions, additions and/or deletions can involve one or more nucleic acids. Alterations can produce conservative or non-conservative amino acid substitutions, deletions and/or additions.

Variants of a prey or a SID® polypeptide encoded by a variant polynucleotide can possess a higher affinity of binding and/or a higher specificity of binding to its protein or polypeptide counterpart, against which it has been initially selected. In another context, variants can also loose their ability to bind to their protein or polypeptide counterpart.

By "anabolic pathway" is meant a reaction or series of reactions in a metabolic pathway that synthesize complex molecules from simpler ones, usually requiring the input of energy. An anabolic pathway is the opposite of a catabolic pathway.

As used herein, a "catabolic pathway" is a series of reactions in a metabolic pathway that break down complex compounds into simpler ones, usually releasing energy in the process. A catabolic pathway is the opposite of an anabolic pathway.

As used herein, "drug metabolism" is meant the study of how drugs are processed and broken down by the body. Drug metabolism can involve the study of enzymes that break down drugs, the study of how different drugs interact within the body and how diet and other ingested compounds affect the way the body processes drugs.

As used herein, "metabolism" means the sum of all of the enzyme-catalyzed reactions in living cells that transform organic molecules.

By "secondary metabolism" is meant pathways producing specialized metabolic products that are not found in every cell.

As used herein, "SID®" means a Selected Interacting Domain and is identified as follows: for each bait polypeptide screened, selected prey polypeptides are compared. Overlapping fragments in the same ORF or CDS define the selected interacting domain.

As used herein the term "PIM®" means a protein-protein interaction map. This map is obtained from data acquired from a number of separate screens using different bait polypeptides and is designed to map out all of the interactions between the polypeptides.

The term "affinity of binding", as used herein, can be defined as the affinity constant Ka when a given SID® polypeptide of the present invention which binds to a polypeptide and is the following mathematical relationship:

$$Ka = \frac{[\text{SID® /polypeptide complex}]}{[\text{free SID®}]\,[\text{free polypeptide}]}$$

wherein [free SID®], [free polypeptide] and [SID® polypeptide complex] consist of the concentrations at equilibrium respectively of the free SID® polypeptide, of the free polypeptide onto which the SID® polypeptide binds and of the complex formed between SID® polypeptide and the polypeptide onto which said SID® polypeptide specifically binds.

The affinity of a SID® polypeptide of the present invention or a variant thereof for its polypeptide counterpart can be assessed for example, on a Biacore™ apparatus marketed by Amersham Pharmacia Biotech Company such as described by Szabo et al *Curr Opin Struct Biol* **5** pgs. 699-705 (1995) and by Edwards and Leartherbarrow, *Anal. Biochem* 246 pgs. 1-6 (1997).

As used herein the phrase "at least the same affinity" with respect to the binding affinity between a SID® polypeptide of the present invention to another polypeptide means that the Ka is identical or can be at least two-fold, at least three fold or at least five fold greater than the Ka value of reference.

As used herein, the term "modulating compound" means a compound that inhibits or stimulates or can act on another protein which can inhibit or stimulate the protein-protein interaction of a complex of two polypeptides or the protein-protein interaction of two polypeptides.

More specifically, the present invention comprises complexes of polypeptides or polynucleotides encoding the polypeptides composed of a bait polypeptide, or a bait polynucleotide encoding a bait polypeptide and a prey polypeptide or a prey polynucleotide encoding a prey polypeptide. The prey polypeptide or prey polynucleotide encoding the prey polypeptide is capable of interacting with a bait polypeptide of interest in various hybrid systems.

As described in the Background of the present invention there are various methods known in the art to identify prey polypeptides that interact with bait polypeptides of interest. These methods, include, but are not limited to, generic two- hybrid systems as described by Fields et al in *Nature*, 340:245-246 (1989) and more specifically in U.S. Patent Nos. 5,283,173, 5,468,614 and 5,667,973, which are hereby incorporated by reference; the reverse two-hybrid system described by Vidal et al, *supra*; the two plus one hybrid method described, for example, in Tirode et al, *supra*; the yeast forward and reverse 'n'-hybrid systems as described in Vidal and Legrain, *supra*; the method described in WO 99/42612; those methods described in Legrain et al *FEBS Letters* 480 pgs. 32-36 (2000) and the like.

The present invention is not limited to the type of method utilized to detect protein-protein interactions and therefore any method known in the art and variants thereof can be used. It is however better to use the method described in WO 99/42612 or WO 00/66722, both references incorporated herein by reference due to the methods' sensitivity, reproducibility and reliability.

Protein-protein interactions can also be detected using complementation assays such as those described by Pelletier et al at http://www.abrf.org/JBT/Articles/JBT0012/jbt0012.html, WO 00/07038 and W098/34120.

Although the above methods are described for applications in the yeast system, the present invention is not limited to detecting protein-protein interactions using yeast, but also includes similar methods that can be used in detecting protein protein interactions in, for example, mammalian systems as described, for example in Takacs et al., Proc. *Nat. Acad. Sci., USA*, **90** (21):10375 (1993) and Vasavada et al., Proc. *Nat. Acad. Sci., USA*, **88** (23):10686-90 (1991), as well as a bacterial two-hybrid system as described in Karimov et al (1998), W099/28746, WO 00/66722 and Legrain et al *FEBS Letters*, **480** pgs. 32-36 (2001).

The above-described methods are limited to the use of yeast, mammalian cells and *Escherichia coli* cells, however the present invention is not limited in this manner. Consequently, mammalian and typically human cells, as well as bacterial, yeast, fungus, insect, nematode and plant cells are encompassed by the present invention and may be transfected by the nucleic acid or recombinant vector as defined herein.

Examples of suitable cells include, but are not limited to, VERO cells, HELA cells such as ATCC No. CCL2, CHO cell lines such as ATCC No. CCL61, COS cells such as COS-7 cells and ATCC No. CRL 1650 cells, W138, BHK, HepG2, 3T3 such as ATCC No. CRL6361, A549, PC12, K562 cells, 293 cells, Sf9 cells such as ATCC No. CRL1711 and Cv1 cells such as ATCC No. CCL70.

Other suitable cells that can be used in the present invention include, but are not limited to, prokaryotic host cells strains such as *Escherichia coli, (e.g.,* strain DH5-α, *Bacillus subtilis, Salmonella typhimurium,* or strains of the genera of *Pseudomonas, Streptomyces* and *Staphylococcus.*

Further suitable cells that can be used in the present invention include yeast cells such as those of *Saccharomyces* such as *Saccharomyces cerevisiae.*

The bait polynucleotide, as well as the prey polynucleotide can be prepared according to the methods known in the art such as those described above in the publications and patents reciting the known method *per se*.

The bait polynucleotide of the present invention is obtained from genomic DNA of *Helicobacter pylori*. The prey polynucleotide is obtained from genomic DNA of *Helicobacter pylori*, variants of genomic DNA of *Helicobacter pylori*, and fragments from the genome or transcriptome of *Helicobacter pylori* ranging from about 20 to 5000. The prey polynucleotide is then selected, sequenced and identified.

A genomic DNA prey library is prepared from the *Helicobacter pylori* and constructed in the specially designed prey vector pP6 as shown in Figure 10 after ligation of suitable linkers such that every genomic DNA insert is fused to a nucleotide sequence in the vector that encodes the transcription activation domain of a reporter gene. Any transcription activation domain can be used in the present invention. Examples include, but are not limited to, Gal4,YP16, B42, His and the like.

Toxic reporter genes, such as CAT^R, CYH2, CYH1, URA3, bacterial and fungi toxins and the like can be used in reverse two-hybrid systems.

The polypeptides encoded by the nucleotide inserts of the genomic DNA prey library thus prepared are termed "prey polypeptides in the context of the presently described selection method of the prey polynucleotides.

The bait polynucleotide can be inserted in bait plasmid as illustrated in Figure 1. The bait polynucleotide insert is fused to a polynucleotide encoding the binding domain of, for example, the Gal4 DNA binding domain and the shuttle expression vector is used to transform cells.

As stated above, any cells can be utilized in transforming the bait and prey polynucleotides of the present invention including mammalian cells, bacterial cells, yeast cells, insect cells and the like.

In an embodiment, the present invention identifies protein-protein interactions in yeast. In using known methods a prey positive clone is identified containing a vector which comprises a nucleic acid insert encoding a prey polypeptide which binds to a bait polypeptide of interest. The method in which protein-protein interactions are identified comprises the following steps:

i) mating at least one first haploid recombinant yeast cell clone from a recombinant yeast cell clone library that has been transformed with a plasmid containing the prey polynucleotide to be assayed with a second haploid recombinant yeast cell clone transformed with a plasmid containing a bait polynucleotide encoding for the bait polypeptide;

ii) cultivating diploid cell clones obtained in step i) on a selective medium; and

iii) selecting recombinant cell clones which grow on the selective medium.

This method may further comprise the step of:

iv) characterizing the prey polynucleotide contained in each recombinant cell clone which is selected in step iii).

In yet another embodiment of the present invention, *in lieu* of yeast, *Escherichia coli is* used in a bacterial two-hybrid system, which encompasses a similar principle to that described above for yeast, but does not involve mating for characterizing the prey polynucleotide.

In yet another embodiment of the present invention, mammalian cells and a method similar to that described above for yeast for characterizing the prey polynucleotide are used.

By performing the yeast, bacterial or mammalian two-hybrid system it is possible to identify for one particular bait an interacting prey polypeptide. The prey polypeptide that has been selected by testing the library of preys in a screen

using the two-hybrid, two plus one hybrid methods and the like, encodes the polypeptide interacting with the protein of interest.

The present invention is also directed, in a general aspect, to a complex of polypeptides, polynucleotides encoding the polypeptides composed of a bait polypeptide or bait polynucleotide encoding the bait polypeptide and a prey polypeptide or prey polynucleotide encoding the prey polypeptide capable of interacting with the bait polypeptide of interest. These complexes are identified in Table 1, as the bait amino acid sequences and the prey amino acid sequences, as well as the bait and prey nucleic acid sequences.

In another aspect, the present invention relates to a complex of polynucleotides consisting of a first polynucleotide, or a fragment thereof, encoding a prey polypeptide that interacts with a bait polypeptide and a second polynucleotide or a fragment thereof. This fragment has at least 20 consecutive nucleotides, but can have between 20 and 5,000 consecutive nucleotides, or between 12 and 10,000 consecutive nucleotides or between 12 and 20,000 consecutive nucleotides.

The polypeptides of column 3 encoded by the polynucleotides of column 2 in Tables 2 and 7 and the polypeptides of column 5 encoded by the polynucleotides of column 3 in Table 8 according to the present invention and the complexes of the two polypeptides encoded by the sets of two polynucleotides also form part of the present invention. In yet another embodiment, the present invention relates to an isolated complex of at least two polypeptides encoded by two polynucleotides wherein said two polypeptides are associated in the complex by affinity binding and are depicted in Table 1 and Table 8.

In yet another embodiment, the present invention relates to an isolated complex comprising at least a polypeptide encoded by an ORF (HP####) of column 1 of Table 1 and a polypeptide encoded by an ORF (HP####) of column 2 of Table 1 and Table 8. The present invention is not limited to these polypeptide complexes alone but also includes the isolated complex of the two polypeptides

in which fragments and/or homologous polypeptides exhibiting at least 95% sequence identity, as well as from 96% sequence identity to 99.999% sequence identity.

Also encompassed in another embodiment of the present invention is an isolated complex in which the SID® polypeptide (see even SEQ ID Nos. from 2 to 3256 in column 3 of Table 2, even SEQ ID Nos. 6590 to 6594 in Table 7 and even SEQ ID Nos. 6596 to 6644 in Table 8.) of the prey polypeptides encoded by uneven SEQ ID Nos. 1 to 3255 in column 2 of Table 2, uneven SEQ ID Nos. 6589 to 6593 in Table 7 and uneven SEQ ID Nos. 6595 to 6643 in Table 8) forming the isolated complex.

Besides the isolated complexes described above, nucleic acids coding for a Selected Interacting Domain (SID®) polypeptide or a variant thereof or any of the nucleic acids set forth in Tables 2, 7 and 8 can be inserted into an expression vector which contains the necessary elements for the transcription and translation of the inserted protein-coding sequence. Such transcription elements include a regulatory region and a promoter. Thus, the nucleic acid which may encode a marker compound of the present invention is operably linked to a promoter in the expression vector. The expression vector may also include a replication origin.

A wide variety of host/expression vector combinations are employed in expressing the nucleic acids of the present invention. Useful expression vectors that can be used include, for example, segments of chromosomal, non-chromosomal and synthetic DNA sequences. Suitable vectors include, but are not limited to, derivatives of SV40 and pcDNA and known bacterial plasmids such as col El, pCR1, pBR322, pMal-C2, pET, pGEX as described by Smith et al (1988), pMB9 and derivatives thereof, plasmids such as RP4, phage DNAs such as the numerous derivatives of phage I such as NM989, as well as other phage DNA such as M13 and filamentous single stranded phage DNA; yeast plasmids such as the 2 micron plasmid or derivatives of the 2 micron plasmid, as well as centomeric and integrative yeast shuttle vectors; vectors useful in eukaryotic cells such as vectors useful in insect or mammalian cells; vectors derived from

combinations of plasmids and phage DNAs, such as plasmids that have been modified to employ phage DNA or the expression control sequences and the like.

For example in a baculovirus expression system, both non-fusion transfer vectors, such as, but not limited to pVL941 (*Bam*HI cloning site Summers, pVL1393 (*Bam*HI, *Sma*I, *Xba*I, *Eco*RI, *Not*I, *Xma*III, *Bgl*II and *Pst*I cloning sites; Invitrogen) pVL1392 (*Bgl*II, *Pst*I, *Not*I, *Xma*III, *Eco*RI, *Xba*II, *Sma*I and *Bam*HI cloning sites; Summers and Invitrogen) and pBlue*Bac*III (*Bam*HI, *Bgl*II, *Pst*I, *Nco*I and *Hind*III cloning sites, with blue/white recombinant screening, Invitrogen), and fusion transfer vectors such as, but not limited to, pAc700 (*Bam*HI and *Kpn*I cloning sites, in which the *Bam*HI recognition site begins with the initiation codon; Summers), pAc701 and pAc70-2 (same as pAc700, with different reading frames), pAc360 (*Bam*HI cloning site 36 base pairs downstream of a polyhedrin initiation codon; Invitrogen (195)) and pBlueBacHisA, B, C (three different reading frames with *Bam*HI, *Bgl*II, *Pst*I, *Nco*I and *Hind*III cloning site, an N-terminal peptide for ProBond purification and blue/white recombinant screening of plaques; Invitrogen (220) can be used.

Mammalian expression vectors contemplated for use in the invention include vectors with inducible promoters, such as the dihydrofolate reductase promoters, any expression vector with a DHFR expression cassette or a DHFR/methotrexate co- amplification vector such as pED *(Pst*I, *Sal*I, *Sba*I, *Sma*I and *Eco*RI cloning sites, with the vector expressing both the cloned gene and DHFR; Kaufman, 1991). Alternatively a glutamine synthetase/methionine sulfoximine co-amplification vector, such as pEE14 *(Hind*III, *Xba*II, *Sma*I, *Sba*I, *Eco*RI and *Bcl*I cloning sites in which the vector expresses glutamine synthetase and the cloned gene; Celltech). A vector that directs episomal expression under the control of the Epstein Barr Virus (EBV) or nuclear antigen (EBNA) can be used such as pREP4 (*Bam*HI, *Sfi*I, *Xho*I, *Not*I, *Nhe*I, *Hind*III, *Nhe*I, *Pvu*II and *Kpn*I cloning sites, constitutive RSV-LTR promoter, hygromycin selectable marker; Invitrogen) pCEP4 (*Bam*HI, *Sfi*I, *Xho*I, *Not*, *Nhe*I, *Hind*III, *Nhe*I, *Pvu*II and *Kpn*I cloning sites, constitutive hCMV immediate early gene promoter, hygromycin selectable marker; Invitrogen), pMEP4 *(Kpn*I, *Pvu*I, *Nhe*I, *Hind*III,

*Not*I, *Xho*I, *Sfi*I, *Bam*HI cloning sites, inducible methallothionein IIa gene promoter, hygromycin selectable marker, Invitrogen), pREP8 (*Bam*HI, *Xho*I, *Not*I, *Hind*III, *Nhe*I and *Kpn*I cloning sites, RSV-LTR promoter, histidinol selectable marker; Invitrogen), pREP9 (*Kpn*I, *Nhe*I, *Hind*III, *Not*I, *Xho*I, *Sfi*I, *Bam*HI cloning sites, RSV-LTR promoter, G418 selectable marker; Invitrogen), and pEBVHis (RSV-LTR promoter, hygromycin selectable marker, N-terminal peptide purifiable via ProBond resin and cleaved by enterokinase; Invitrogen).

Selectable mammalian expression vectors for use in the invention include, but are not limited to, pRc/CMV (*Hind*III, *Bst*XI, *Not*I, *Sba*I and *Apa*I cloning sites, G418 selection, Invitrogen), pRc/RSV (*Hind*III, *Spe*I, *Bst*XI, *Not*I, *Xba*I cloning sites, G418 selection, Invitrogen) and the like. Vaccinia virus mammalian expression vectors (see, for example Kaufman 1991) that can be used in the present invention include, but are not limited to, pSC11 (*Sma*I cloning site, TK- and β-gal selection), pMJ601 (*Sal*I, *Sma*I, *Afl*II, *Nar*I, *Bsp*MII, *Bam*HI, *Apa*I, *Nhe*I, *Sac*II, *Kpn*I and *Hind*III cloning sites; TK- and β-gal selection), pTKgptF1S (*Eco*RI, *Pst*I, *Sal*I, *Acc*I, *Hind*II, *Sba*I, *Bam*HI and *Hpa* cloning sites, TK or XPRT selection) and the like.

Yeast expression systems that can also be used in the present invention include, but are not limited to, the non-fusion pYES2 vector (*Xba*I, *Sph*I, *Sho*I, *Not*I, *Gst*XI, *Eco*RI, *Bst*XI, *Bam*HI, *Sac*I, *Kpn*I and *Hind*III cloning sites, Invitrogen), the fusion pYESHisA, B, C (*Xba*I, *Sph*I, ShoI, *Not*I, *Bst*XI, *Eco*RI, *Bam*HI, *Sac*I, *Kpn*I and *Hind*III cloning sites, N-terminal peptide purified with ProBond resin and cleaved with enterokinase; Invitrogen), pRS vectors and the like.

Consequently, mammalian and typically human cells, as well as bacterial, yeast, fungi, insect, nematode and plant cells an used in the present invention and may be transfected by the nucleic acid or recombinant vector as defined herein.

Examples of suitable cells include, but are not limited to, VERO cells, HELA cells such as ATCC No. CCL2, CHO cell lines such as ATCC No. CCL61,

COS cells such as COS-7 cells and ATCC No. CRL 1650 cells, W138, BHK, HepG2, 3T3 such as ATCC No. CRL6361, A549, PC12, K562 cells, 293 cells, Sf9 cells such as ATCC No. CRL1711 and Cv1 cells such as ATCC No. CCL70.

Other suitable cells that can be used in the present invention include, but are not limited to, prokaryotic host cells strains such as *Escherichia coli*, (e.g., strain DH5-α), *Bacillus subtilis*, *Salmonella typhimurium*, or strains of the genera of *Pseudomonas*, *Streptomyces* and *Staphylococcus*.

Further suitable cells that can be used in the present invention include yeast cells such as those of *Saccharomyces* such as *Saccharomyces cerevisiae*.

Besides the specific isolated complexes, as described above, the present invention relates to and also encompasses SID® polynucleotides. As explained above, for each bait polypeptide, several prey polypeptides may be identified by comparing and selecting the intersection of every isolated fragment that are included in the same polypeptide, as set forth, in Example 5. Thus the SID® polynucleotides of the present invention are represented by the nucleic acid sequences of uneven SEQ ID Nos. 1 to 3255 in column 2 of Table 2, uneven SEQ ID Nos. 6589 to 6593 in Table 7 and uneven SEQ ID Nos. 6595 to 6643 in Table 8 encoding the SID® polypeptides of even SEQ ID Nos. 2 to 3256. of Table 2, the even SEQ ID Nos. 6590 to 6594 in Table 7 and the even SEQ ID Nos. 6596 to 6644 in Table 8.

The present invention is not limited to the SID® nucleic acid sequences as described in the above paragraph, but also includes fragments of these sequences having at least 6 consecutive nucleic acids, between 6 and 5,000 consecutive nucleic acids and between 6 and 10,000 consecutive nucleic acids and between 6 and 20,000 consecutive nucleic acids, as well as variants thereof. The fragments or variants of the SID® sequences possess at least the same affinity of binding to its protein or polypeptide counterpart, against which it has been initially selected. Moreover this variant and/or fragments of the SID® sequences alternatively can have between 95% and 99.999% sequence identity to its protein or polypeptide counterpart.

According to the present invention the variants can be created by known mutagenesis techniques either *in vitro* or *in vivo*. Such a variant can be created such that it has altered binding characteristics with respect to the target protein and more specifically that the variant binds the target sequence with either higher or lower affinity.

Polynucleotides that are complementary to the above sequences which include the polynucleotides of the SID®'s, their fragments, variants and those that have specific sequence identity are also included in the present invention.

The polynucleotide encoding the SID® polypeptide, fragment or variant thereof can also be inserted into recombinant vectors which are described in detail above.

The present invention also relates to a composition comprising the above-mentioned recombinant vectors containing the SID® polypeptides in Tables 2, 7 and 8, fragments or variants thereof, as well as recombinant host cells transformed by the vectors. The recombinant host cells that can be used in the present invention were discussed in greater detail above.

The compositions comprising the recombinant vectors can contain physiological acceptable carriers such as diluents, adjuvants, excipients and any vehicle in which this composition can be delivered therapeutically and can include, but are not limited to sterile liquids such as water and oils.

In yet another embodiment, the present invention relates to a method of selecting modulating compounds, as well as the modulating molecules or compounds themselves which may be used in a pharmaceutical composition. These modulating compounds may act as a cofactor, as an inhibitor, as antibodies, as tags, as a competitive inhibitor, as an activator or alternatively have agonistic or antagonistic activity on the protein-protein interactions.

The activity of the modulating compound does not necessarily, for example, have to be 100% activation or inhibition. Indeed, even partial activation or inhibition can be achieved that is of pharmaceutical interest.

The modulating compound can be selected according to a method which comprises:

(a) cultivating a recombinant host cell with a modulating compound on a selective medium and a reporter gene the expression of which is toxic for said recombinant host cell wherein said recombinant host cell is transformed with two vectors:

(i) wherein said first vector comprises a polynucleotide encoding a first hybrid polypeptide having a DNA binding domain;

(ii) wherein said second vector comprises a polynucleotide encoding a second hybrid polypeptide having a transcriptional activating domain that activates said toxic reporter gene when the first and second hybrid polypeptides interact;

(b) selecting said modulating compound which inhibits or permits the growth of said recombinant host cell.

Thus, the present invention relates to a modulating compound that inhibits the protein-protein interactions of a complex of two polypeptides of Table 1 and Table 8. The present invention also relates to a modulating compound that activates the protein-protein interactions of a complex of two polypeptides of Table 1 and Table 8.

In yet another embodiment, the present invention relates to a method of selecting a modulating compound, which modulating compound inhibits the interactions of two polypeptides of Table 1. This method comprises:

(a) cultivating a recombinant host cell with a modulating compound on a selective medium and a reporter gene the expression of which is toxic for said recombinant host cell wherein said recombinant host cell is transformed with two vectors:

(i) wherein said first vector comprises a polynucleotide encoding a first hybrid polypeptide having a first domain of an enzyme;

(ii) wherein said second vector comprises a polynucleotide encoding a second hybrid polypeptide having an enzymatic transcriptional activating domain that activates said toxic reporter gene when the first and second hybrid polypeptides interact;

(b) selecting said modulating compound which inhibits or permits the growth of said recombinant host cell.

In the two methods described above any toxic reporter gene can be utilized including those reporter genes that can be used for negative selection including the URA3 gene, the CYH1 gene, the CYH2 gene and the like.

In yet another embodiment, the present invention provides a kit for screening a modulating compound. This kit comprises a recombinant host cell which comprises a reporter gene the expression of which is toxic for the recombinant host cell. The host cell is transformed with two vectors. The first vector comprises a polynucleotide encoding a first hybrid polypeptide having a DNA binding domain; and a second vector comprises a polynucleotide encoding a second hybrid polypeptide having a transcriptional activating domain that activates said toxic reporter gene when the first and second hybrid polypeptides interact.

In yet another embodiment a kit is provided for screening a modulating compound by providing a recombinant host cell, as described in the paragraph above, but instead of a DNA binding domain, the first vector comprises a first hybrid polypeptide containing a first domain of a protein. "The second vector comprises a second polypeptide containing a second part of a complementary domain of a protein that activates the toxic reporter gene when the first and second hybrid polypeptides interact.

In the selection methods described above, the activating domain can be p42 Gal 4, YP16 (HSV) and the DNA-binding domain can be derived from Gal4 or Lex

A. The protein or enzyme can be adenylate cyclase, guanylate cyclase, DHFR and the like.

SID® in Tables 2, 7 and 8 may be used as modulating compounds.

In yet another embodiment, the present invention relates to a pharmaceutical composition comprising the modulating compounds for preventing or treating ulcers in a human or animal, most preferably in a mammal.

This pharmaceutical composition comprises a pharmaceutically acceptable amount of the modulating compound. The pharmaceutically acceptable amount can be estimated from cell culture assays. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes or encompasses a concentration point or range having the desired effect in an *in vitro* system. This information can thus be used to accurately determine the doses in other mammals, including humans and animals.

The therapeutically effective dose refers to that amount of the compound that results in amelioration of symptoms in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or in experimental animals. For example, the LD50 (the dose lethal to 50% of the population) as well as the ED50 (the dose therapeutically effective in 50% of the population) can be determined using methods known in the art. The dose ratio between toxic and therapeutic effects is the therapeutic index which can be expressed as the ratio between LD50 and ED50 compounds that exhibit high therapeutic indexes.

The data obtained from the cell culture and animal studies can be used in formulating a range of dosage of such compounds which lies preferably within a range of circulating concentrations that include the ED50 with little or no toxicity.

The pharmaceutical composition can be administered via any route such as locally, orally, systemically, intravenously, intramuscularly, mucosally, using a patch and can be encapsulated in liposomes, microparticles, microcapsules, and

the like. The pharmaceutical composition can be embedded in liposomes or even encapsulated.

Any pharmaceutically acceptable carrier or adjuvant can be used in the pharmaceutical composition. The modulating compound will be preferably in a soluble form combined with a pharmaceutically acceptable carrier. The techniques for formulating and administering these compounds can be found in *"Remington's Pharmaceutical Sciences"* Mack Publication Co., Easton, PA, latest edition.

The mode of administration optimum dosages and galenic forms can be determined by the criteria known in the art taken into account the seriousness of the general condition of the mammal, the tolerance of the treatment and the side effects.

The present invention also relates to a method of treating or preventing ulcers in a human or mammal in need of such treatment. This method comprises administering to a mammal in need of such treatment a pharmaceutically effective amount of a modulating compound which binds to a targeted bacterial protein. In a preferred embodiment, the modulating compound is a polynucleotide which may be placed under the control of a regulatory sequence which is functional in the mammal or human.

In yet another embodiment, the present invention relates to a pharmaceutical composition comprising a SID® polypeptide, a fragment or variant thereof. The SID® polypeptide, fragment or variant thereof can be used in a pharmaceutical composition provided that it is endowed with highly specific binding properties to a bait polypeptide of interest.

The original properties of the SID® polypeptide or variants thereof interfere with the naturally occurring interaction between a first protein and a second protein within the cells of the organism. Thus, the SID® polypeptide binds specifically to either the first polypeptide or the second polypeptide.

Therefore, the SID® polypeptides of the present invention or variants thereof interfere with protein-protein interactions of *Helicotacter pylori* proteins or between *Helicobacter pylori* proteins and mammal, for example, human proteins.

Thus, the present invention relates to a pharmaceutical composition comprising a pharmaceutically acceptable amount of a SID® polypeptide or variant thereof, provided that the variant has the above-mentioned two characteristics; i.e., that it is endowed with highly specific binding properties to a bait polypeptide of interest and is devoid of biological activity of the naturally occurring protein.

In yet another embodiment, the present invention relates to a pharmaceutical composition comprising a pharmaceutically effective amount of a polynucleotide encoding a SID® polypeptide or a variant thereof wherein the polynucleotide is placed under the control of an appropriate regulatory sequence. Appropriate regulatory sequences that are used are polynucleotide sequences derived from promoter elements and the like.

Polynucleotides that can be used in the pharmaceutical composition of the present invention include the nucleotide sequence: of uneven SEQ ID Nos. 1 to 3255 in column 2 of Table 2, uneven SEQ ID Nos. 6589 to 6593 in Table 7 and uneven SEQ ID Nos. 6595 to 6643 in Table 8.

Besides the SID® polypeptides and polynucleotides, the pharmaceutical composition of the present invention can also include a recombinant expression vector comprising the polynucleotide encoding the SID® polypeptide, fragment or variant thereof.

The above described pharmaceutical compositions can be administered by any route such as orally, systemically, intravenously, intramuscularly, intradermally, mucosally, encapsulated, using a patch and the like. Any pharmaceutically acceptable carrier or adjuvant can be used in this pharmaceutical composition.

The SID® polypeptides as active ingredients will be preferably in a soluble form combined with a pharmaceutically acceptable carrier. The techniques for formulating and administering these compounds can be found in *"Remington's Pharmaceutical Sciences" supra.*

The amount of pharmaceutically acceptable SID® polypeptides can be determined as described above for the modulating compounds using cell culture and animal models.

Such compounds can be used in a pharmaceutical composition to treat or prevent ulcer.

Thus, the present invention also relates to a method of preventing or treating ulcer in a mammal said method comprising the steps of administering to a mammal in need of such treatment a pharmaceutically effective amount of:

(1) a SID® polypeptide of even SEQ ID Nos. 2 to 3256 in column 3 of Table 2, even SEQ ID Nos. 6590 to 6594 in Table 7 and even SEQ ID Nos. 6596 to 6644 in Table 8 or variants thereof which binds to a targeted mammalian or typically human protein; or

(2) or SID® polynucleotide encoding a SID® polypeptide of uneven SEQ ID Nos. 1 to 3255 in column 2 of Table 2, uneven SEQ ID Nos. 6589 to 6593 in Table 7 and uneven SEQ ID Nos. 6595 to 6643 in Table 8 or variants or fragments thereof wherein said polynucleotide is placed under the control of a regulatory sequence which is functional in said mammal; or

(3) a recombinant expression vector comprising a polynucleotide encoding a SID® polypeptide which binds to a bacterial protein.

In another embodiment the present invention nucleic acids comprising a sequence which encodes the SID® proteins of Table 2 Table 7 and Table 8 and/or functional derivatives thereof are administered to modulate complex of Table 1 and Table 8 by way of gene therapy. Any of the methodologies relating to gene therapy available within the art may be

used in the practice of the present invention such as those described by Goldspiel et al *Clin. Pharm.* ***12*** pgs. 488-505 (1993).

Delivery of the therapeutic nucleic acid into a patient may be direct *in vivo* gene therapy (i.e., the patient is directly exposed to the nucleic acid or nucleic acid containing vector) or indirect *ex vivo* gene therapy (i.e., cells are first transformed with the nucleic acid *in vitro* and then transplanted into the patient).

For example for *in vivo* gene therapy, an expression vector containing the nucleic acid is administered in such a manner that it becomes intracellular; i.e., by infection using a defective or attenuated retroviral or other viral vectors as described, for example in U.S. Patent 4,980,286 or by Robbins et al, Pharmacol. *Ther.*, ***80*** No. 1 pgs. 35-47 (1998).

The various retroviral vectors that are known in the art are such as those described in Miller et al, *Meth. Enzymol.* ***217*** pgs. 581-599 (1993) which have been modified to delete those retroviral sequences which are not required for packaging of the viral genome and subsequent integration into host cell DNA. Also adenoviral vectors can be used which are advantageous due to their ability to infect non-dividing cells and such high-capacity adenoviral vectors are described in Kochanek, *Human Gene Therapy,* ***10***, pgs. 2451-2459 (1999). Chimeric viral vectors that can be used are those described by Reynolds et al, *Molecular Medecine Today*, pgs. 25 -31 (1999). Hybrid vectors can also be used and are described by Jacoby et al, *Gene Therapy,* ***4***, pgs. 1282-1283 (1997).

Direct injection of naked DNA or through the use of microparticle bombardment (e.g., Gene Gun®; Biolistic, Dupont) or by coating it with lipids can also be used in gene therapy. Cell-surface receptor s/transfecting agents or through encapsulation in liposomes, microparticles or microcapsules or by administering the nucleic acid in linkage to a peptide which is known to enter the nucleus or by administering it in linkage to a ligand predisposed to receptor-mediated endocytosis (See, Wu & Wu, *J. Biol. Chem.,* ***262*** pgs. 4429-4432; (1987)) can be used to target cell types which specifically express the receptors of interest.

In another embodiment a nucleic acid ligand compound may be produced in which the ligand comprises a fusogenic viral peptide designed so as to disrupt endosomes, thus allowing the nucleic acid to avoid subsequent lysosomal degradation. The nucleic acid may be targeted *in vivo* for cell specific endocytosis and expression by targeting a specific receptor such as that described in W092/06180, W093/14188 and WO 93/20221. Alternatively the nucleic acid may be introduced intracellularly and incorporated within the host cell genome for expression by homologous recombination. See, Zijlstra et al, Nature, 342, pgs. 435-428 (1989).

In *ex vivo* gene a gene is transferred into cells *in vitro* using tissue culture and the cells are delivered to the patient by various methods such as injecting subcutaneously, application of the cells into a skin graft and the intravenous injection of recombinant blood cells such as hematopoietic stem or progenitor cells.

Cells into which a nucleic acid can be introduced for the purposes of gene therapy include, for example, epithelial cells, endothelial cells, keratinocytes, fibroblasts, muscle cells, hepatocytes and blood cells. The blood cells that can be used include, for example, T-lymphocytes, B-lymphocytes, monocytes, macrophages, neutrophils, eosinophils, megakaryocytes, granulocytes, hematopoietic cells or progenitor cells and the like.

In yet another embodiment the present invention relates to protein chips or protein microarrays. It is well known in the art that microarrays can contain more than 10,000 spots of a protein that can be robotically deposited on a surface of a glass slide or nylon filter. The proteins attach covalently to the slide surface, yet retain their ability to interact with other proteins or small molecules in solution. In some instances the protein samples can be made to adhere to glass slides by coating the slides with an aldehyde-containing reagent that attaches to primary amines. A process for creating microarrays is described, for example by MacBeath and Schreiber in *Science,* Volume 289, Number 5485, pgs, 1760-1763 (2000) or Service, *Science,* Vol, 289, Number 5485 pg. 1673 (2000). An

apparatus for controlling, dispensing and measuring small quantities of fluid is described, for example, in U.S. Patent No. 6,112,605.

The present invention also provides a record of protein-protein interactions, PIM®'s, SID®'s and any data encompassed in the following Tables. It will be appreciated that this record can be provided in paper or electronic or digital form.

In yet another embodiment, the present invention relates to the classification of *H. pylori* within functional categories such as genes essential or non essential for viability using the general method described in Figure 18. In this regard, two exhaustive libraries of *H. pylori* ORFs were constructed in *E.coli* The first library contained every *H. pylori* (strain 26695) cloned individually (Library I) while the second one (Library II), contained these ORFs disrupted by a transposable element. These two ordered libraries are valuable tools for a large project of systematic inactivation of every ORF of the *H. pylori* genome. They were used to develop a strategy to search at the genomic scale for genes essential for the viability of the bacterium grown *in vitro*. The inactivation strategy was applied to a series of 138 ORFs that were selected on two different criteria. Ninety six of them were previously shown to encode either proteins involved in protein-protein interaction in the two-hybrid yeast assay (Rain et al, 2001), and 42 encode *H. pylori* specific protein with no known function. The screening procedure led to the identification of 40 Putative Essential Genes (PEGs), of which 15 were shown to be true essential genes. The combination of both essentiality as well as the identification of interacting domains might serve as a direct pathway for the design of active compounds capable of inhibiting protein-protein interactions and possibly bacterial growth was the reasoning used behind this analysis.

Construction and validation of the two ordered libraries

Library I consisted of an individual and ordered (*hp0001* to *hp1590*) bank of every putative ORF according to the information provided on the Webb TIGR site in 1997. Based on the 1590 ORFs identified at that time, 5' and 3'-oligonucleotides were synthesized with the characteristics described in Example 10 and in Table 3. Each ORF was cloned into a pILL570-• derivative plasmid, 3.6

kb in length. This derivative plasmid corresponds to pILL570 (Labigne et al, 1992) in which DNA from the *Hind*III site of the polylinker to the *Ava*I site (position 1425 of the pBR322 bone) has been excluded by reverse PCR using pILL570 as a template and 570-1 plus 570-2 as primers (Table 4). Each ORF was cloned in such way that the 5'-end of the gene (including the ATG) was inserted immediately downstream the three transcriptional and translational stops of pILL570 (Labigne et al, 1992) to prevent toxicity of the recombinant proteins in *E. coli*. The library consists of 17 times 96-well plates (plate I.1 to plate XVII.1). The recombinant plasmids were transformed into DH5α *E. coli* cells harboring the pTCA plasmid, a plasmid that confers resistance to tetracycline, encodes the Tn*3* transposase and is immune to Tn*3* (Seifert et al, 1986). The presence of the two compatible plasmids pILL570-HP000X plus pTCA was checked by plasmid extraction and gel electrophoresis on individual isolated tetracycline, spectinomycin, kanamycin resistant clones. In addition, using primers 570-3 and 570-4 (corresponding to the bounderies of the cloning site on pILL570-•) the adequation between the size of the cloned PCR product and that of the corresponding ORF was confirmed. Library I consists of all the putative ORFs described on the TIGR Webb site in 1997 with the exclusion of 40 ORFs (*hp01, 10, 46, 56, 94, 160, 223, 264, 289, 293, 399, 415, 435, 440, 453, 159, 464, 465, 468, 547, 607, 722, 790, 814, 846, 876, 884, 898, 968, 1007, 1069, 1205, 1248, 1304, 1358, 1394, 1452, 1460, 1497, 1511*) for which either the initial gene amplification or the final cloning failed.

Library II consists of the random insertion of Tn*3*-Km into each of the recombinant pILL570-••*hp000X* plasmid. Process of disruption has been designed to generate multiple independent transposon insertions for each cloned ORF. Tn*3*-Km was shown to preferentially map into the *H. pylori* inserts due to both the intrinsic properties of Tn*3* that transposes into AT rich DNA region and the requirement of maintaining intact replicative function and spectinomycin modifying enzyme (*aadA*). The efficiency of the whole procedure was checked for five plates. For those 5 plates, the resulting kanamycin transconjugants of 96 independent cloned ORFs were kept individually and as pools of plasmids. The 96 disrupted recombinant plasmids were extracted together and used as a

template for individual PCR assays using as primers the 38bp-R (Table 4) of the inverted repeats of Tn3 transposon and each of the 5'-end specific oligonucleotide used for the cloning of the ORFs. Figure 17 illustrates the results obtained for the disruption of the ORFs *hp0099* to *hp0198* (Plate II.2). This figure exemplifies first that multiple insertions of the transposon took place, second that for the majority trransposon insertion occurred at a distance ranging between 100 to 600 bp from the 5'-end of the ORF, a distance compatible with the promotion of gene replacement by allelic recombination.

Screening of Putative Essential Genes of *H. pylori* within a series of 138 selected ORFs

The ordered library of disrupted *H. pylori* ORFs in *E. coli* was used for the genomic screening of putative essential genes (PEGs), and the screening of a subset of 96 individual selected ORFs of *H. pylori* strain 26695 (Table 5) as genes encoding proteins demonstrating homodimeric or heterodimeric protein-protein interactions [Rain, 2001 #14], and the screening of a subset of 42 ORFs encoding *H. pylori* specific predicted protein with no known function (Table 6). Included in the set of 138 ORFs were ORFs used as controls: ORFs known to be essential [*groES (hp0011); holB* (*hp1231*), *dnaA* (*hp1529*], or ORFs known to be non-essential for viability *in vitro*: *ureI*, (*hp0071*); *rdxA* (*hp0954*); *ggt* (*hp1118*) representive of various sizes. For the 138 ORFs, individual Tn3-Km disrupted recombinant plasmids were extracted from Library II, and used to transform *H. pylori* strain HAS141 (Janvier et al, 1999). This isolate was chosen for the initial screening due to its high natural transformation competency, two logs above that of the sequenced strain 26695, and its ability to colonize mouse stomach (Janvier et al, 1999). Kanamycin transformants were obtained for all but the *hspA* gene as expected, and 40 of the 138 tested ORFs, namely *hp0061, 175, 377, 419, 553, 650, 739, 862, 928, 990, 1012, 1014, 1074, 1230, 1245, 1263, 1493* for the first series (Table 5), and *hp0130, 231, 271, 358, 394, 659, 697, 699, 721, 726, 746, 838, 935, 947, 953, 973, 1023, 1028, 1039, 1053, 1085, 1265, 1568* (Table 6) which thus can be designated as Putative Essential Genes. Due to the presence of the two large terminal inverted repeats (38 bp in size) within Tn3-Km

which quickly reannealed during the amplification procedure, gene amplification of the disrupted allele using the respective 5' and 3' oligonucleotides of a given ORF often led to the production of a PCR product with a size similar to that of the parental allele, and thus was not helpful to confirm gene replacement of the parental by the disrupted allele. Thus, six individual kanamycin resistant transformants obtained for the 78 knock-out (KO) genes were controlled by gene amplification using the 5' and the 3' oligonucleotides of the KO-ORF respectively in pairs with the 38 bp (Table 4) of the inverted repeat of the transposon. Criteria for allelic replacement were that the sum of the size of the two PCR products be identical to that of the KO-ORF. The final identification of the disrupted ORF and of the site of insertion was done by sequencing one of the two PCR products. For ORFs with a size over 700 bp two or three different transposon insertions mapping in the middle of the ORF were commonly observed among the six analysed transformants. For a few ORFs, the 5' and the 3' oligonucleotides initially designed from the sequence of strain 26695, were unable to amplify the chromosome of the HAS141 parental strain due to intrinsic polymorphism. In these cases, the 5' and 3' oligos from ORFn+1 or ORFn-1 together with the 38 bp were used. Unexpectedly, a few kanamycin resistant transformants were obtained for *hp1231* (*holB*), and *hp1529* (*dnaA*). For these genes the insertion transposon mapped at the very 3' end of the genes, in contrast to the other genes where the transposable element mapped at any place but at a minimal distance of 300 bp of either end of the genes. These observations indicate that for some of the genes, known to be essential, the kanamycin resistant gene could be rescued, without allelic replacement, by integration of the whole plasmid via a single crossing-over. To estimate the frequency of this event, 260 preparations of chromosomal DNA from individual 210) or pooled (50) kanamycin resistant *H. pylori* transformants were spotted on nitrocellulose, denatured and hybridized with a probe consisting of the pILL570-• vector. Of 260 DNA preparations tested, chromosomal DNA extracted from individual transformants of the disrupted *hp1231*(*holB*), *1514* (unknown) and *1529* (*dnaA*) as well as some (and not all) of the individual clones of *hp0224* and *hp0822* did hybridize with part of the vector confirming that a single crossing-over took place. These results underlined that

the *Tn3*-Km strategy is a powerful approach to be used as a first screen, at a genomic scale, for the identification of PEGs. The definitive assignment of a non-essential status cannot be done exclusively on the presence of kanamycin transformants, but had to be confirmed and tested for absence of vector DNA within the chromosome of transformants by hybridization.

Essentiality and functional analysis of 15 of the 40 ORFs identified as Putative Essential Genes

According to an in-house definition, PEGs consist of ORFs that did not lead to the isolation of kanamycin transformants following the transformation of the parental isolate HAS141 with a pool of disrupted ORF. However, absence of kanamycin transformants cannot be directly associated with the identification of a true essential gene. Several explanations might account for the lack of knock out kanamycin resistant mutant, such as (i) absence of the specific gene in the tested strain, (ii) polar effect of the transposon on an essential ORF located downstream of the PEG, a property associated with miniTn3-Km (Skouloubris et al, 1998) (iii) experimental failure due to the small size of the ORF and of the bordering sequences required for allelic exchange. Thus, to confirm the essential character of a PEG (i.e., essentiality) additional experiments were achieved. For each of the 40 PEGs identified through the initial screening in strain HAS141 (Tables 5 & 6), the following criteria were used to confirm their essentiality. First, the distribution of the different ORFs of the genome by DNA/DNA hybridizations within clinical isolates to confirm their presence not only in the transformed isolate (HAS141), but also in every clinical isolate tested was studied. Second, a non-polar mutation was introduced into the PEG cloned in *E. coli* following the approach depicted in the Materials and Methods section, performed by reverse PCR from plasmids of Library I. The recombinant plasmids consist, in each case, of the 0.9 kb-Kanamycin promoter-less cassette flanked respectively upstream and downstream with the 300 first and the 300 last nucleotides of the ORF to knock-out. In addition, the cassette carries a ribosome binding site and a start codon (ATG) in phase with the hundred 3'-terminal codons of the KO gene allowing the translation of the end of the gene to overcome any

transcriptional/translational coupling effect. The resulting constructed plasmid was transformed in four different *H. pylori* genetic backgrounds: strains HAS141, N6, X47-2An and the sequenced strain 26695. Taking into account these criteria (Figure 17), the following conclusions were drawn relative to the 40 initially identified PEGs.

Strain HAS141 lacks ORFs *hp0990* and *hp1074*. These absences accounted for the lack of mutant for these two ORFs, whereas the other 38 PEGs were found to be present in each of the 37 tested clinical isolates (Salama et al, 2000) and manuscript in preparation). The fact that no HAS141 clones resistant to kanamycin were obtained when transformed with the derivative plasmids containing the disrupted *hp0990* and *hp1074* ORFs confirmed that in the absence of portable homologous region integration of the plasmid does not occur. These two ORFs were not studied further since a non ubiquitous ORF has low probability to be essential and even if so, would not be a good candidate for *H. pylori* specific therapeutical drug design.

Non polar kanamycin mutants were unambiguously obtained in HAS141, N6, X47-2An as well as 26695 for ORFs *hp0061, 419, 553, 650, 1263*, and *1493* of the first series of genes encoding interacting protein (Table 5), and *hp0130, 271, 358, 697, 699, 721, 726, 746, 838, 935, 947, 953, 973, 1023, 1028, 1039, 1053* (Table 6). Gene replacement of the parental allele by the deleted and disrupted allele was confirmed by testing the chromosome of the mutant for the disappearance of a PCR product with a size identical to that of the parental ORF and its replacement by the expected size (1.5 kb PCR product: 300 + 900+ 300 bp). When by chance, the parental ORF had a size ranging around 1.5 kb, gene replacement was confirmed by restricting the 1.5 kb PCR product with *SmaI* to release the 0.9 kb-Km non-polar cassette.

For 15 of the 38 ubiquitous PEGs, we were unable to obtain kanamycin transformants when introducing a non polar mutation in N6, X47-2An, 26695, although the genes were found to be present in every isolates (ubiquitous). One or two very rare clones were obtained for some of the genes exclusively in HAS141. All these kanamycin resistant HAS141 transformants were positive

when tested by hybridization with the labeled pILL570-• vector used as a probe, again attesting for the presence of rare but possible single crossing-over seen under strong selective pressure. These 15 genes (*hp0175, hp0231, hp0377, hp0394, hp0659, hp0739, hp0862, hp0928, hp1012, hp1014, hp1085, hp1230, hp1245, hp1265, hp1568*) can thus be definitively recognized as genes essential for the viability of *H. pylori in vitro*. Among these 15 genes, 9 are known to be genes that encode proteins involved in protein-protein interactions, and 6 were selected as encoding *H. pylori* specific protein without known function. They encode proteins with properties that will be discussed and classified in regards to their potential as putative therapeutical targets.

Search for interacting proteins when using the 6 ubiquitous *H. pylori* specific essential genes as a bait in the two hybrid assay

The *hp0231, hp0394, hp0659, hp1085, hp1265,* and *hp1568* ORFs were cloned in the pB6vector and used as individual bait for the identification of interacting proteins (Table 7).

Proteomic screens conducted for HP0231, HP1085, HP1568 did not provide data allowing assignation of a putative function and did not reveal homodimeric interaction underline the usefulness of the protein as a possible therapeutic target. Those genes remain ubiquitous, essential, *H. pylori* specific and without known function. In contrast, HP0394, HP0659 and HP1265 gave positive screens (Table 7).

Classification of twelve of the 15 ubiquitous essential genes.

The combination of both essentiality as well as the identification of interacting domains might serve as a direct pathway for the design of active compounds capable of inhibiting protein-protein interactions and possibly bacterial growth. Of the 15 ubiquitous essential genes identified by the procedure, 12 were shown to be involved in protein-protein interaction and could be classified in different categories with regards to their potential as putative therapeutic targets.

The first category consists of ubiquitous ORFs encoding proteins with heterodimeric protein-protein interactions in the two-hybrid assay where both partners are playing an essential role for the viability of *H. pylori* and at least one of the two partners is *H. pylori* specific. Four of the twelve essential ORFs answer these criteria *hp0394*, *hp0862*, *hp1230*, and *hp0659* which encode proteins with no known or putative function. Although initially annotated as *H. pylori* specific, the recent publication of the *Campylobacter* genome demonstrates for some of them the existence of homologues in this closely phylogenetically related bacteria, but no homologues have been identified in the other bacterial genomes so far sequenced.

The *hp1230* gene encodes a protein that has been recognized via the two hybrid assay as an homodimeric protein which interacts with the predicted chromosomal replication initiator protein, DnaA, encoded by *hp1529*. The proteomic screen allowed the identification of a specific domain of interaction (SID) lying between AA31 and AA180 of HP1230 (SID1230) and a SID of 87 AA within the N-terminal domain of HP1529 or DnaA (AA12 to AA99) (SID1529) (Rain *et al.*, 2001). To examine whether the HP1230/HP1529 interaction was specific and could serve as a target for screening of lead compounds with bactericidal activity, the oligonucleotide encoding SID1529 was randomly mutagenized, and selected, through the two-hybrid system, mutated *sid* that abolished the specific HP1230/HP1529 interaction. This allowed the demonstration of the fact that isoleucine 58 and Lysine 61 were involved in the HP1230/HP1529 interaction since a double mutant I58F/K61I within *sid1529* abolished this interaction. This mutagenesis random procedure also led to the isolation of a mutated SID1529* (V53L) which confers a superbinder phenotype to HP1529 in the two-hybrid system. When overexpressed in the three-hybrid system (Tirode *et al.*, 1997) under the control of the regulated Met25 promoter in vector p3H1 (Colland *et al.*, 2001) (Figure 3), the superbinder SID1529* completely inhibits the HP1230/HP1529 interaction. Thus, oligopeptide – PNQLLCTTITAKYG-(SEQ ID No. 6588) or overlapping or combinatory derivatives have some potential as lead compounds to inhibit *H. pylori* growth.

Another example of this category of interest is the *hp0862* gene. The two-hybrid screen procedure revealed interactions between the HP0862 gene product and the C-terminal domain (AA100 to AA191) of the thymidylate kinase (HP1474), an essential enzyme responsible for the first phosphorylation step in the conversion of deoxythymidine 5'-monophosphate to deoxythymidine 5'-diphosphate for the final production of dTTP. As for HP1230, the actual function of the HP0862 encoding gene is unknown, but its essential character and interaction with a known essential enzyme (Tmk) might orient further functional analysis, and encouraged the definition of more precise domain of interactions between the TmK protein and this new specific interacting protein.

The *hp394* and *hp659* might also enter this category. They both encode protein essential specific to *H. pylori*. The *hp394* encodes a protein whose C-terminal domain (last 76 AA) interacts with the ß subunit of the acetylcoenzyme A carboxylase transferase (HP950). This enzyme has an essential function in membrane lipid synthesis and catalyses the formation of malonyl-CoA, the first intermediate for fatty acid synthesis . The protein encoded by *hp0659* interacts with a putative outer membrane protein HP0655 with no known function. The essential property of *hp0655* has not yet been tested and thus the classification of the *hp0659* gene in this first category has to be confirmed.

- The second category includes genes essential for *H. pylori* encoding predicted proteins with known functions. The characteristic of this category is that these genes essential in *H. pylori* have not been reported to be lethal in other organisms. Thus targeting the proteins encoded by these genes might be relevant of a selective drug design specifically directed against *H. pylori*. Three of the twelve essential genes displayed this characteristic: *hp0377* and *hp0175*, they both encode periplasmic proteins. *Hp*0377 encodes a protein partially homologue to DsbC a thiol-disulfide interchange periplasmic protein involved in disulfide bond formation (Zapun *et al.*, 1995). So far, no other genes encoding Dsb-like proteins have been identified in the genome of *H. pylori* whereas in *E. coli,* as many as five proteins are required for disulfide bond formation in the periplasm. In *E. coli, dsbC* encodes a stable homodimer with both protein disulfide

isomerase and chaperone activities (McCarthy *et al.*, 2000), and is not essential probably due to redundancy of the function. In contrast, in *H. pylori* we demonstrate the essential character of that gene. The HP0377 product interacts in the two-hybrid assay with the last 100 C-terminal amino acids (SID) of the homodimeric secreted cysteine containing protein encoded by *hp0224*. The HP0224 product is a methionine sulfoxide reductase homologue (MsrA). In *E. coli*, MsrA plays a role in response to oxidative damage by reducing the methionine sulfoxide residues (Moskovitz *et al.*, 1995), and directly, or indirectly contributes to the maintenance of adhesins (Wizemann *et al.*, 1996). The *H. pylori* MsrA homologue (HP0224) was identified as one of the major antigens released into the extracellular space (Cao *et al.*, 1998). Thus, the HP0377 encoding DsbC is a good *H. pylori* specific target candidate because of the essential character unique to the bacterial species, and the accessibility of the protein within the periplasmic space.

The *hp0175* gene is another representative of this category. The gene encodes a predicted peptidyl-propyl cis-trans isomerase, an enzyme that accelerates protein folding by catalyzing the cis-trans isomerization of proline imidic peptide bonds in oligopeptides. Two-hybrid screening identified HP0608, a *H. pylori* specific protein of unknown function, as interacting partner with HP0175. The essential character of the gene, documented in this study, as for HP0377, has not been reported for other microorganisms and appears unique to the species.

Finally, *hp1265* is a very unique and specific target for *H. pylori*. This essential gene is part of a large cluster of 14 genes (*hp1260* to *hp1273*) among which 12 encode homologues of 12 of the 14 subunits of the NADH oxidoreductase complex of *E. coli*. In *E. coli* inactivation of the genes encoding the different subunits (*nuo*) is not lethal for the bacterium. In *H. pylori* inactivation of *hp1263* is not lethal, but inactivation of one of the two *H. pylori* specific subunit encoding gene is lethal, indicating that the HP1265 subunit which should be the subunit involved in the NADH binding has very unique properties deserving further functional investigations.

- A third category, consists in ORFs that encoded conserved hypothetical proteins, distributed in almost all the sequenced bacterial genomes, for which no function has been assigned, nor the essentiality assessed. We report here on the essential character of two of them the *hp0739* gene which encodes a protein that interacts with another conserved hypothetical but non essential protein (HP810), and *hp1012*. In the two sequenced *H. pylori* genomes (26695 and J99), the *hp0739* gene is flanked by two genes involved in the biosynthesis of the peptidoglycan, and its involvement in this biosynthetic pathway remains to be explored. The *hp1012* gene encoded a protein, which has similitude with some metallo-proteases, however the function of this specific protease is unknown. These two proteins with large distribution spectrum, so far unknown as essential genes, represent new putative targets.

- Finally, a fourth category, includes ORFs with known functions that were previously shown to be essential in other microorganisms. The present work allows us to extend this property to the *H. pylori* species, and reinforces their value as putative targets with large spectrum. The *hp0928* gene is one of those; it was selected through the two-hybrid screen as an homodimer. By similarity, the assigned function was that of GTP cyclohydrolase (*folE)* involved in the first step of the biosynthetic pathway of tetrahydrofolate, the structure of which was shown to be a homodecameric complex form of two pentamers. Both, the inability to succeed in knocking out that gene, and its oligomeric structure supports the assigned function. The two-hybrid screening procedure delineates a domain of interaction consisting of 133 amino-acids between the FolE subunits. This domain might serve as a therapeutical target for the screening of lead compounds with large bacterial spectrum. The *hp1014* gene (*hdhA*) encodes an NAD^+-dependent oxidoreductase belonging to the short-chain dehydrogenase/reductase (SDR 1 family). The enzyme is known to require a tetrameric form to be active in *E. coli* (Yoshimoto *et al.*, 1991) which was compatible with the homodimeric interaction observed in the two hybrid assay. The major interest of these two ORFs resides in their well known enzymatic activity, which provides a direct way to screen lead compounds capable of

abolishing the enzymatic activity through the disruption of the oligomeric interactions. The present work also classifies *hp1245* as an essential gene of this category. This gene encodes the SSB protein, the single strand binding protein, involved in DNA replication, recombination and DNA repair, and such observation is confirmatory of work previously done in other model microorganisms. HP1245 was found to interact with HP0650, a non–essential protein of unknown function, but also significantly with HP0661, a predicted ribonuclease H involved in DNA replication, a finding consistent with the HP1245 function.

In order to fully illustrate the present invention and advantages thereof, the following specific examples are given, it being understood that the same are intended only as illustrative and in nowise limitative.

EXAMPLES

Example 1: Preparation of a *Helicobacter pylori* genomic collection

1.A. Fragment collection preparation arid transformation in E. coli

1.B. Collection transformation in *Saccharomyces cerevisiae*

1.C. Construction of bait plasmid

Example 2: Screening the collection with the two-hybrid in yeast system

2.A. The mating protocol

2.B. The X-Gal overlay assay

2.C. The luminometry assay

Example 3: Identification of positive clones

3.A. PCR on yeast colonies

3.B. Plasmid rescue from yeast by electroporation

Example 4: Detection of protein-protein interaction

Example 5: Identification of SID®

Example 6: Screening of modulating agent

Example 7: Gene therapy Example using SID® polypeptides

Example 8: Making of polyclonal and monoclonal antibodies

Example 9: Classification of genes of H. pylori

Example 10: Study of the interaction between two essential genes, HP1230-HP1529, by random mutagenesis.

Medium composition and standard protocols are available in Maniatis *et al.*.

Example 1: Preparation of a *Helicobacter pylori* genomic collection

1.A. Collection preparation and transformation in Escherichia coli

1.A.1. Fragment of genomic DNA preparation

The *Helicobacter pylori* genomic DNA is fragmented in a nebulizer (GATC) for 1 minute, precipitated and resuspended in water.

The obtained nebulized genomic DNA is successively treated with Mung Bean Nuclease (Biolabs) (30 minutes at 30°C), T4 DNA polymerase (Biolabs) (10 minutes at 37°C) and Klenow enzyme (Pharmacia) (10 minutes at room temperature and 1 hour at 16°C).

DNA is then extracted, precipitated and resuspended in water.

1.A.2. Ligation of linkers to blunt-ended genomic DNA

Oligonucleotide PL160 (5' end phosphorylated) 1 µg /µl and PL159 2µg/µl.

Sequence of the oligo PL160 : 5'-ATCCCGGACGAAGGCC-3' (SEQ ID NO. 3257)

Sequence of the oligo PL159 : 5'-GGCCTTCGTCCGG-3' (SEQ ID NO. 3258)

Linkers were preincubated (5 minutes at 95°C, 10 minutes at 68°C, 15 minutes at 42°C) then cooled down at room temperature and ligated with genomic DNA inserts at 4°C overnight.

Linkers were further removed on a separation column (Chromaspin TE 400, Clontech), according to the manufacturers protocol.

1.A.3. Vector preparation

pACTIIst is successively digest with *Bam*HI restriction enzyme (Biolabs) for 1 hour at 37°C, dephosphorylated with Calf Intestine Phosphatase (CIP) (Biolabs) and filled in with dGTP using Vent DNA polymerase (exo-) (Biolabs), extracted, precipitated and resuspended in water.

1.A.4. Ligation between vector and insert of genomic DNA

The prepared vector is ligated overnight at 15°C with the genomic blunt ended DNA described in section 2 using T4 DNA ligase (Biolabs). The DNA is then precipitated and resuspended in water.

1.A.5. Library transformation in Escherichia coli

Transform DNA from section 1.A.4 into Electromax DH10B electrocompetent cells (Gibco BRL) with a Cell Porator apparatus (Gibe BRL). Add 1 ml SOC medium and incubate transformed cells at 37°C for 1 hour. Add 9 ml volume of SOC medium per tube and plate on LB+ampicillin medium. Scrape colonies with liquid LB medium. Aliquot and freeze at -80°C.

The obtained collection of recombinant cell clones was named HGXBHP1 (CNCM N° I-2181).

1.B. Collection transformation in *Saccharomyces cerevisiae*

The *Saccheramyces cerevisiae* strain (Y187 (MATα GalΔ Gal80Δ ade2-101 His3 Leu2-3, -112 Trp1-901 Ura3-52 URA3::UASGAL1-LacZ Met)) transformed with the HGXBHP1 *H. pylori* genomic DNA library.

The plasmid DNA contained in *E. coli* were extracted (Qiagen) from aliquoted *E. coli* frozen cells (1.A.5.).

Grow Saccharomyces cerevisiae yeast Y187 in YPGlu.

Yeast transformation is performed according to standard protocol (Giest et al. Yeast, 11, 355-360, 1995) using yeast carrier DNA (Clontech). This experiment leads to 10^4 to 5 10^4 cells/µg DNA. Spread 2 10^4 cells on DO-Leu medium per plates. Aliquot and freeze at -80°C.

1. C. Construction of bait plasmid

The genomic amplification of the ORF is obtained by PCR using the Pfu proofreading *Taq* polymerase (Stratagene) and 200 ng of genomic DNA as the template. PCR primers are chosen in regions flanking the ORF.

Set up the PCR program as followed:

94°	45"	
94°	45"	x 30 cycles
48°	45"	
72°	6'	
72°	10'	
15°	∞	

Check amplification on agarose gel.

Purify PCR fragments with Qiaquick column (Qiagen) according to the manufacturer's protocol.

Digest purified PCR fragments with adequate restriction enzymes.

Purify PCR fragments with Qiaquick column (Qiagen) according to the manufacturer's protocol.

Ligate digested PCR fragments into an adequately digested and dephosphorylated bait vector (pAS2ΔΔ) according to standard protocol (Maniatis et al.).

Transform into competent bacterial cells. Grow cells, extract DNA and sequence plasmid.

Example 2: Screening the collection with the two-hybrid in yeast system

2.A. The mating protocol

We have chosen the mating two-hybrid in yeast system (first described by Legrain et al., Nature Genetics, 1997, vol. 16, 277-282, *Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens)* for its advantages but we could also screen the *Helicobacter* pylori collection in the classical two-hybrid system as described in Fields *et al,* or in a yeast reverse two-hybrid system.

The mating procedure allows a direct selection on selective plates because the two fusion proteins are already produced in the parental cells. No replica plating is required.

This protocol is written for the use of the library transformed into the Y187 strain. Before mating, transform *S. cerevisiae* (CG 1945 strain (MATa Gal4-542 Gal180-538 ade2-101 His3*200 Leu2-3,-112 Trp1-901 Ura3-52 Lys2-801 URA3::GAL4 17mers (X3)-CyC1TATA-LacZ LYS2::GAL1UAS-GAL1TATA-HIS3 CYH^R)) according to step 1.B and spread on DO-Trp medium.

Day 1, morning : preculture

Preculture of Y187 cells carrying the bait plasmid obtained at step 1.C. in 20 ml DO-Trp medium. Grow at 30°C with vigorous agitation.

Day 1, late afternoon : culture

Measure OD_{600}nm of the DO-Trp preculture of Y187 cells carrying the bait plasmid preculture. The OD_{600}nm must lie between 0.1 and 0.5 in order to correspond to a linear measurement.

Inoculate 150 ml DO-Trp at OD_{600}nm 0.006/ml, grow overnight at 30°C with vigorous agitation.

Day 2 : mating

Medium and plates

5 YPGlu plates

50 ml tube with 30 ml DO-Leu-Trp-His

100 ml flask with 20 ml of YPGlu

75 DO-Leu-Trp-His plates

2 DO-Leu plates 2 DO-Trp plates

2 DO-Leu-Trp plates

Measure OD_{600}nm of the DO-Trp culture. It should be around 1.

For the mating, you must use twice as many bait cells as library cells. To get a good mating efficiency, you must collect the cells at 10^8 cells per cm^2.

Estimate the amount of bait culture (in ml) that makes up 80 OD_{600}nm units for the mating with the prokaryote library.

Thaw a vial containing the HGXYHP1 library slowly on ice. Add the contents of the vial to 20 ml YPGlu. Let those cells recover at 30°C, under gentle agitation for 10 minutes.

Mating

Put the 80 OD_{600}nm units of bait culture into a 250 ml flask.

Add the HGXYHP1 library culture to the bait culture. Transfer the mixture of diploids into 50 ml sterile tubes. Centrifuge, discard the supernatant and resuspend in YPGlu medium.

Distribute cells in 400 µl samples in YPGlu plates with glass beads. Spread cells by shaking the plates.

Incubate plates cells-up at 30°C for 4h30min.

Collection of mated cells

Wash and rinse plates and spread collected cells on DO-Leu-Trp-His+Tet plates.

Day 4

Selection of clones able to grow on DO-Leu-Trp-His+-Tetracyclin : this medium allows us to isolate diploid clones presenting an interaction.

Count the His+ colonies on control plates.

The number of His+ cell clones will define which protocol is to be processed:

Upon 20 $X10^6$ His+ colonies

- if the number of His+ cell clones > 285 : then process overlay and then luminometry protocols on blue colonies (2.B and 2.C).

- If the number of His+ cell clones < 285 : process luminometry protocol (2.C).

- The following step leads to the selection of the strongest interaction.

2.B. The X-Gal overlay assay

The X-Gal overlay assay is performed directly on the selective medium plates after scoring the number of His+ colonies.

Material

Set up a waterbath. The water temperature should be 50°C.

- 0.5 M Na_2HP0_4 pH 7.5.
- 1.2% Bacto-agar.
- 2% X-Gal in DMF.
- Overlay mixture : 0.25 M Na_2HPO_4 pH7.5, 0.5% agar, 0.1% SDS, 7% DMF (LABOSI), 0.04% X-Gal (ICN). For each plate, 10 ml overlay mixture are needed.
- DO-Leu-Trp-His plates.
- Sterile toothpicks.

Experiment

Temperature of the overlay mix should be between 45 and 50°C.

Pour the overlay-mix over the plates in portions of 10 ml.

Collect them when the top layer is settled.

Incubate plates overlay-up at 30°C. Note the time.

Check for blue colonies regularly. If no blue colony appears, wait for overnight incubation. Mark with a pen and number the positives.

Streak the positives colonies on fresh DO-Leu-Trp-His plates with a sterile toothpick.

2. C. The luminometry assay

Grow His+ colonies overnight at 30°C in microtiter plates containing DO-Leu-Trp-His+Tetracyclin medium with shaking. The day after, dilute 15 times overnight culture into a new microtiter plate containing the same medium. Incubate 5 hours at 30°C with shaking. Dilute samples 5 times and read OD_{600}nm. Dilute again to obtain between 10,000 and 75,000 yeast cells/well in 100 µl final volume.

Per well, add 76 µl of One Step Yeast Lysis Buffer (Tropix), 20 µl Sapphirell Enhancer (Tropix), 4 µl Galacton Star (Tropix), incubate 40 minutes at 30°C. Measure the β-Gal read-out (L) using a Luminometer (Trilux, Wallach).

Calculate value of OD_{900nm}xL and select interacting preys having the highest values.

At this step of the protocol, we have isolated diploid cell clones presenting interactions. The next step is now to identify polypeptides involved in the selected interactions.

Example 3: Identification of positive clones

3.A. PCR on yeast colonies

Introduction

PCR amplification of fragments of plasmid DNA directly on yeast colonies is a quick and efficient procedure to identify sequences cloned into this plasmid. It is directly derived from a published protocol (Wang H. et al., Analytical Biochemestry, 237, 145-146, 1996). However, it is not a standardized protocol : in our hands it varies from strain to strain, it is dependent of experimental conditions (number of cells, *Taq* polymerase source, etc). This protocol should be optimized to specific local conditions.

Materials

- For 1 well, PCR mix composition is:

  32.5 µl water,

  5 µl 10X PCR buffer (Pharmacia),

  1 µl dNTP 10 mM,

  0.5 µl *Taq* polymerase (5µl /µl) (Pharmacia),

  0.5 µl oligonucleotide ABS1 10 pmole/µl:

5'-GCGTTTGGAATCACTACAGG-3' (SEQ ID NO. 3259)

0.5 µl oligonucleotide ABS2 10 pmole/µl:

5'-CACGATGCACGTTGAAGTG-3' (SEQ ID NO. 3260)

- 1 N NaOH.

Experiment

Grow positive colonies overnight at 30°C on a 96 well cell culture cluster (Costar), containing 150 µl DO-Leu-Trp-His+Tetracyclin with shaking. Resuspend culture and transfer immediately 100 µl on a Thermowell 96 (Costar).

Centrifuge 5 minutes at 4,000 rpm at room temperature.

Remove supernatant. Dispense 5 µl NaOH in each well, shake 1 minute.

Place the Thermowell in the thermocycler (GeneAmp 9700, Perkin Elmer) 5 minutes at 99.9°C and then 10 minutes at 4°C.

In each well, add PCR mix, shake well.

The PCR program was set up as followed :

94°C	3 minutes		
94°C	30 secondes		
53°C	1 minute 30 secondes	x 35 cycles	
72°C	3 minutes		
72°C	5 minutes		
15°C	∞		

The quality, the quantity and the length of the PCR fragment was checked on an agarose gel. The length of the cloned fragment was the estimated length of the PCR fragment minus 300 base pairs that corresponded to the amplified flanking plasmid sequences.

3.B. Plasmids rescue from yeast by electroporation

Introduction

The previous protocol of PCR on yeast cell may not be successful, in such a case, plasmids from yeast by electroporation can be rescued. This experiment allows the recovery of prey plasmids from yeast cells by transformation of *E. coli* with a yeast cellular extract. The prey plasmid can then be amplified and the cloned fragment can be sequenced.

Material

Plasmid rescue

Glass beads 425-600 µm (Sigma)

Phenol/chloroform (1/1) premixed with isoamyl alcohol (Amresco)

Extraction buffer: 2% Triton X100, 1 % SDS, 100 mM NaCl, 10 mM TrisHCl pH 8.0, 1 mM EDTA pH 8.0.

Mix ethanol/NH_4Ac : 6 volumes ethanol with 7.5 M NH_4 Acetate, 70% Ethanol and yeast cells in patches on plates.

Electroporation

SOC medium

M9 medium

Selective plates: M9-Leu+Ampicillin

2 mm electroporation cuvettes (Eurogentech)

Experiment

Plasmid rescue

Prepare cell patch on DO-Leu-Trp-His with cell culture of section 2.C.

Scrape the cell of each patch into an Eppendorf tube, add 300 µl of glass beads in each tube, then, add 200 µl extraction buffer and add 200 µl phenol:chloroform: isoamyl alcohol (25:24:1).

Centrifuge tubes 10 minutes at 15,000 rpm.

Transfer 180 µl supernatant to a sterile Eppendorf tube and add to each 500 µl ethanol/NH_4Ac, vortex.

Centrifuge tubes 15 minutes, 15,000 rpm at 4°C

Wash pellet with 200 µl 70% ethanol, remove ethanol and dry pellet.

Resuspend pellet in 10 µl water. Store extracts at -20°C.

Electroporation

Material : Electrocompetent MC1066 cells prepared according to standard protocols (Maniatis).

Add 1 µl of yeast plasmid DNA-extract to pre-chilled Eppendorf tube, and keep on ice.

Mix 1 µl plasmid yeast DNA-extract sample, add 20 µl electrocompetent cells and transfer in a cold electroporation cuvette.

Set the Biorad electroporator on 200 ohms resistance, 25 µF capacity; 2.5 kVolts. Place cuvette in the cuvette holder and electroporate.

Add 1 ml SOC into the cuvette and transfer the cell-mix into a sterile Eppendorf tube. Let cells recover for 30 minutes at 37°C, spin the cells down 1 minute, 4,000x g and pour off supernatant. Keep about 100 µl medium and use it to resuspend the cells and spread them on selective plates (e.g., M9-Leu plates).

Incubate plates for 36 hours at 37°C.

Grow one colony and extract plasmids. Check presence and size of insert through enzymatic digestion and agarose gel. Sequence insert.

Example 4: Protein-protein interaction

For each bait, the previously protocol leads to the identification of prey polynucleotide sequences. In order to identify a protein-protein interaction, we need to characterize the obtained prey polypeptide sequence regarding the *Helicobacter* pylori genome.

This may be accomplish with a software program names blastwun (available on the Internet site of the University of Washington http://bioweb.pasteur.fr/seqanal/interfaces/blastwu.html, this is a development version of software for gene and protein identification through similarity searches of protein and nucleotide sequence databases).

Blastwun program compares prey polynucleotide insert sequence (rescued from prey plasmid) with whole *Helicobacter pylori* genome (available on NCBI web site: http://www.ncbi.nlm.nih.gov under GenBank accession number AE000511). This comparison leads to prey polynucleotide localizations in the *H. Pylori* genome, each localization having a score depending on the homology of sequence. For each prey polynucleotide, we consider the localization with the highest score and, if the insert sequence is included in and is in phase with an Open Reading Frame, we can identify one prey polypeptide interacting with one bait polypeptide.

Helicobacter pylori ORF's sequences are available on the World-Wide Web site of The Institute for Genomic Research (TIGR) at http://www.tigr.org/tdb/mdb/hpdb/hpdb.html.

This web page allows several requests concerning *Helicobacter pylori*'s genome, in particular, its ORF sequence. To get the sequences of specific ORF's, click on the ewindow named "HP#" and click search. This operation leads to a new web page presenting nucleic and peptide sequence of the specific ORF.

See Table 1 : Protein-protein interactions in *Helicobacter pylori*.

Example 5: Identification of SID®

Experiment results in step 4. sequences of each prey fragment encoding for an interacting prey polypeptide.

By comparing and selecting the intersection of every isolated fragments that are included in the same polypeptide, we define the Selected Interacting Domain (SID®) see Figure 15.

See results in Table 2.

Example 6: Screening of modulating agent

Select one specific interaction.

Transform a permeabilized yeast cell with plasmids containing bait polypeptide and prey polypeptide of the specific interaction.

Plate a top agar containing transformed permeabilized yeast cells on square boxes (that already contains agarose gel).

Apply by spotting the compounds to test on top agar as soon as it is solidified. Incubate overnight at 30°C.

Analyse results : select lead compounds that prevent transformed permeabilized yeast cells from growing.

Example 7: Gene therapy Example using SID® polypeptides

An expression vector containing the SID® polynucleotide is made in the manner described in U.S. Patent 4,980,286. It is then administered to patients to treat *H. pylori* infections.

Example 8: Making of polyclonal and monoclonal antibodies

The protein-protein complex of Table 1 was injected into mice and polyclonal and monoclonal antibodies were made following the procedure set forth in Sambrook et al *supra*.

More specifically mice are immunized with an immunogen comprising complexes conjugated to keyhole limpet hemocyanin using glutaraldehyde or

EDC as is well known in the art. The complexes can also be stabilized by crosslinking as described in WO 00/37483. The immunogen is then mixed with an adjuvant. Each mouse receives four injections of 10 μg to 100 μg of immunogen, and after the fourth injection, blood samples are taken from the mice to determine if the serum contains antibodies to the immunogen. Serum titer is determined by ELISA or RIA. Mice with sera indicating the presence of antibody to the immunogen are selected for hybridoma production.

Spleens are removed from immune mice and single-cell suspension is prepared (Harlow et al 1988). Cell fusions are performed essentially as described by Kohler et al. Briefly, P365.3 myeloma cells (ATTC Rockville, Md) or NS-1 myeloma cells are fused with spleen cells using polyethylene glycol as described by Harlow et al. Cells are plated at a density of 2×10^5 cells/well in 96-well tissue culture plates. Individual wells are examined for growth and the supernatants of wells with growth are tested for the presence of Table 1 complex-specific antibodies by ELISA or RIA using the Table 1 complex as a target protein. Cells in positive wells are expanded and subcloned to establish and confirm monoclonality.

Clones with the desired specificities are expanded and grown as ascites in mice or in a hollow fiber system to produce sufficient quantities of antibodies for characterization and assay development. Antibodies are tested for binding to bait polypeptide (from column 1 of Table 1) alone or to prey polypeptide (from column 2 of Table 1) alone, to determine which are specific for the Table 1 complex as opposed to those that bind to the individual proteins.

Monoclonal antibodies against each of the complexes set forth in Table 1 are prepared in a similar manner by mixing specified proteins together, immunizing an animal, fusing spleen cells with myeloma cells and isolating clones which produce antibodies specific for he protein complex, but not for individual proteins.

Example 9: Classification of genes of *H. pylori* within function categories at the genomic scale using 2 exhaustive libraries in *E. coli*

1. Bacterial strains, growth and storage conditions.

Escherichia coli strains DH5α (BRL) HB101 (Boyer and Roulland-Dussoix, 1969) and NS2114 (Rif^R) (Seifert *et al.*, 1986) were used as hosts for plasmid cloning and disruption experiments and were grown at 37°C in L-broth (10 g of tryptone, 5 g of yeast extract and 5 g of NaCl per liter, pH 7.0) or on L-agar plates (1.5% agar) at 37°C. Antibiotics were used at the following final concentrations (μg/ml) unless indicated in the text: spectinomycin:100 (Upjohn Laboratories, Paris, France), tetracycline: 8 (Sigma Chemicals, Saint-Quentin Fallavier, France), kanamycin: 25 (Serva, Frankfurt, Germany), rifampicine:100 μg (Sigma Chemicals). Independent recombinant *E. coli* were saved by storing up to 96 clones individually in 96-well microtitre plates; clones were inoculated into L-broth supplemented with 8 μg/ml tetracycline, 100 μg/ml spectinomycin and 7% DMSO (Sigma) and stored at -80°C. *H. pylori* strain 26695 (Tomb *et al.*, 1997), HAS141 (Janvier *et al.*, 1999), N6 (Ferrero *et al.*, 1992), X47-2an (GUY et al., 1999) were routinely cultured on 10% horse blood agar medium (Blood Agar Base no. 2; Oxoid, Lyon, France) or in Brucella broth supplemented with 10 % Fecal Calf Serum (Gibco). Solid and liquid media contained supplements at the following final concentrations: 10 μg vancomycin (Dakota Pharmaceuticals, Creteil, France), 2.5 IU polymyxin (Pfizer Laboratories, Orsay, France), 5 μg trimethoprim (Sigma) and 4 μg amphotericin B (Bristol-Myers Squibb, Paris, France)/ml. Plates were incubated at 37°C under microaerobic conditions in an anaerobic jar with a carbon dioxide generator (CampyGen, Oxoid) without catalyst. *H. pylori* that had undergone chromosomal allelic exchange were selected on medium supplemented with 25 μg kanamycin.

Production of amplicons corresponding to each of the 1590 ORFs originally identified by TIGR in strain 26695: 1621 pairs of forward and reverse oligonucleotides targeting the 1590 ORFs of the genome of strain 26695, as assigned on the Webb site in 1997, were designed and synthesized by Eurogentec (Bel S.A., Seraing, Belgium). Pairs (sense and antisense, Table 3) of oligonucleotides were designed in order to allow full length amplification of each of the ORFs, with the exception of ORFs with a size over 3 kb that were split into

two or three PCR products. Every forward and reverse 21 bp-oligonucleotide was tagged respectively at its 5'extremity with a CAUCAUCAU (Seq ID No.3261) and a CUACUACUA (SEQ ID No. 3262) sequence. Genomic DNA of *Hp* strain 26695, prepared from cesium chloride extraction (Labigne-Roussel *et al.*, 1988), was used as a template for PCR. Seventeen times 96 amplicons were produced by polymerase chain reactions (PCR) using a PCR-express thermal cycler through 40 cycles consisting of a denaturation step of 94°C for 2 mn, a primer annealing step of 50°C for 30 sec. and an extension step at 72°C for 2 mn, under a 96-well format. The amplicons were controlled for size and quality (single band) on agarose gel.

Systematic UDG cloning of the ORFs of *H. pylori* strain 26695

The cloning of the 96 amplicons was performed using the ligation-independent method described by Rashtchian (Rashtchian, 1995). First the linear pILL570-• derivative vector was prepared by gene amplification using #570-1 and #570-2 (Table 4) as primers and the pILL570 plasmid (Labigne *et al.*, 1992) as a template. Three microliters of individual HP0001 to HP1590 PCR products were mixed together with 2 µl of pILL570-• derivative vector (75 ng), 14 µl of 1XPCR buffer, 1 µl of uracil DNA glycosylase (UDG) in a 2000-µl 96-well disposable plate. The plates were incubated at 37°C for 30 mn allowing the enzymatic reaction as well as the hybridization between protruding and complementary extension hybridization to occur. Competent DH5α cells (100 µl) harboring the pTCA plasmid (Seifert *et al.*, 1986) were added to each well, and the 96-well plate was further incubated for 45 mn on ice. One ml of prewarmed L-broth was added to each well, and the plate was then incubated for 90 mn at 37°C. Finally, a selective antibiotic cocktail containing spectinomycin, tetracycline was added to each well to positively select and enrich in pILL570-• derivative recombinant plasmid transformed DH5α(pTCA) cells; plates were then incubated for another 13 hours at 37°C under agitation. Individual transformant colonies were isolated by spotting 10 µl of liquid culture from each well on square agar plates containing tetracycline and spectinomycin using a 96-well inoculator designed to deliver a 10

µl liquid volume; cloning of the PCR product was confirmed by mini-preparation recombinant plasmid restricted with *ClaI-AvaI*. They were stored in DMSO (7%) at –80°C under a 96-well format as "library I" consisting of plate I.1 to plate XVII.1.

Systematic disruption of the recombinant plasmids by transposon mutagenesis in *E. coli*. Transposon mutagenesis of individual *E. coli* clones was performed using the mini-Tn*3*-Km transposon as previously described by Jenks *et al.* (Jenks *et al.*, in press). All manipulations were performed in a 96-well format and four independent transposon mutageneses were carried out in parallel so as to saturate the mutagenesis disrupting process with independent events. Briefly, the stored microtitre plates containing the individual *E. coli* DH5α clones that harbored the pTCA plus the recombinant pILL570-• derivative plasmids were thawed and used to inoculate fresh plates. Plasmid pILL553 harboring the mini-Tn*3*-Km transposon (Seifert *et al.*, 1986) (Labigne, 1997) (a low copy auto-transferable plasmid pOX38 derivative) was transferred into these *E. coli* DH5α clones by conjugation. Transconjugates harboring all three plasmids (recombinant pILL570-• derivative, pTCA and pILL553), were selected by spotting 10 µl of the mating mixture on L-agar containing 25 µg/ml kanamycin, 8 µg/ml tetracycline and 100 µg/ml spectinomycin. Cointegrates were transferred by conjugation into *E. coli* NS2114SmRif carrying the *cre* gene. Positive selection of resolved forms of the cointegrates was obtained by growth on L-agar containing 100 µg/ml rifampicin (Sigma), 625 µg/ml kanamycin, 625 µg/ml spectinomycin and 625 µg/ml streptomycin. A pool of each TnKm disrupted recombinant plasmids was stored as individual stock of disrupted ORFs at -80°C in DMSO (7%), designated "library II" consisting of plate I.2 to plate XVII.2.

DNA preparation and standard molecular biology techniques. The alkaline lysis procedure (Sambrook *et al.*, 1989) was used for small-scale preparation and MIDI Qiagen (Qiagen, Courtaboeuf, France) columns were used for large-scale plasmid preparation. Whole cell genomic DNA from individual generated *H. pylori* mutants was extracted using the QIAamp Tissue Kit (Qiagen) according to the

manufacturer's instructions. Standard procedures for DNA/DNA hybridization were prepared on nitrocellulose membranes (Schleicher and Schuell,) according to the procedure of Sambrook et al. (Sambrook *et al.*, 1989). They were hybridized under standard conditions with [α-^{32}P]-deoxyribonucleotide probes labeled by random priming using the MegaPrime DNA system (Amersham) according to the manufacturer's instructions.

Transformation of *H. pylori*. *H. pylori* strains were naturally transformed with circular plasmid DNA (~ 2 µg per transformation). Briefly, bacteria were inoculated as 1 cm patches and grown for 5 h before addition of 10 µl supercoiled plasmid DNA. Each disrupted plasmid consisting either of a pool of disrupted plasmids when originating from library II or of a single recombinant plasmid for the non polar mutation construction was added to 4 independently prepared patches of *H. pylori*. After further incubation for 18 h, the bacteria from each individual patch were harvested and plated directly onto a single plate of selective medium (kanamycin, 25 µg/ml). Six individual kanamycin transformants were then subcultured. Chromosomal DNA was extracted using QiAmp kit extraction, and the constructed mutant characterized by several PCR controls and/or hybridization as described in the result section.

Introduction of a non polar mutation in a selected ORF (HP000X)

For a given cloned ORF (HP000X) present in Library I, the recombinant plasmid was prepared and used as a template for reverse PCR performed with two oligonucleotides. Oligonucleotide HP000X-1 consists of 24 nucleotides complementary to an intragenic sequence located 300 bp downstream the 5'-end of the gene, and HP000X-2 consists of 24 nucleotides complementary to an intragenic sequence located 300 bp upstream of the 3'-end of the gene. All the HP00X-1 and HP000X-2 oligos were tagged respectively at their 5'-end with a CGGGGTACC (SEQ ID No. 3261) (*Kpn*I) sequence and at their 3'-end with a CGCGGATCC (SEQ ID No. 3262) (*Bam*HI) sequence (Table 4). Following reverse amplification with high fidelity *Taq* polymerase (Boerhinger) from the given cloned ORF, 5 µl of PCR product were restricted with *Dpn*I in order to

eliminate the template molecule, then restricted with *Kpn*I and *Bam*HI and directly ligated to the promoterless non polar *Kpn*I-*Bam*HI kanamycin cassette (0.9 kb) previously described (Skouloubris *et al.*, 1998). Transformants were selected on spectinomycin containing plates. Recombinant plasmids were then purified, controlled, and introduced by natural transformation into *H. pylori* cells.

Example 10 – Study of the interaction between two essential genes, HP1230-HP1529, by random mutagenesis.

1. Preparation of mutagenized SID1529 collection

A. Collection preparation and transformation in *Escherichia coli*

i. A Random mutagenesis of SID1529 by PCR

Mutagenized SID1529 was obtained by PCR using the Taq polymerase (Stratagene) and 200 ng of *Helicobacter pylori* genomic DNA and the following oligonucleotides :

5'-ATTTGCGGCCGCAATCTTGGCGCTAGTCAAACAA-3' (SEQ ID No. 3263)

5'-CCGGGATCCTCAAGATTGGGCGTTAATTTGGAT-3' (SEQ ID No. 3264)

The PCR program was set up as follows :

94° 1'

45° 1' x 30 cycles

75° 1'

The PCR conditions were as follows :

KCl (50 mM), Tris pH 9.0 (10 mM), $MgCl_2$ (7 mM), $MnCl_2$ (0.2 mM)

The amplification was checked by agarose gel electrophoresis.

B. Digestion of mutagenized SID1529 and pP7-centro vector

The PCR fragments were purified with Qiaquick column (Qiagen) according to the manufacturer's protocol and digested (*NotI-BamHI*). The vector (pP7-centro) (see, Figure 20) was digested (*NotI-BamHI*) and dephosphorylated according to standard protocol (Sambrook *et al.*).

C. Library transformation in *Escherichia coli*

The method followed in this section is the same as the one described in Example 1.A.5 above.

2. Collection transformation in Saccharomyces cerevisiae

The method followed in this section is the same as the one described in Example 1.B above.

3. Construction of HP1230 bait plasmid

The genomic amplification of the HP1230 ORF was obtained by PCR using the Pfu proofreading Taq polymerase (Stratagene) and 200 ng of *Helicobacter pylori* genomic DNA as template.

The PCR program was set up as follows:

94°	45"		
94°	45"		
48°	45"	x 30 cycles	
72°	6'		
72°	10'		

The amplification was checked by agarose gel electrophoresis.

The PCR fragments were purified with Qiaquick column (Qiagen) according to the manufacturer's protocol. The digested PCR fragments were ligated into an adequately digested (BamHI-PstI) and dephosphorylated bait vector (pB1) according to standard protocol (Sambrook *et al.*) and were transformed into competent bacterial cells. The cells were grown , the DNA extracted and the plasmid was sequenced.

4. Screening the mutagenized SID1529 collection with the HP1230 bait protein using two-hybrid in yeast

The method followed in this section is the same as the one described in Example 2 above, with the exception that DO Leu-Trp-His + Tet plates were replaced by DOLeu-Trp-His + Tet + 40 mM 3-AT plates.

5. Identification of positive clones

The method followed in this section is the same as the one described in Example 3 above.

6. The SID 1529* (V53L) inhibits the interaction between HP1529 and HP1230

In the three-hybrid system, the HP1529 protein is expressed fused to the GAL4 Activation Domain (AD) in the pP6 plasmid, whereas HP1230 is introduced in the p3H1 vector in fusion with the DNA-binding domain (DBD) of GAL4. In addition, this vector contains the Met25 promoter which allow expression of a third partner in medium lacking methionine. After transformation of Y187 and CG1945 yeast by the pP6-HP1529 and p3H1-HP1230 vectors, respectively, both strains were mated. The resulting diploid strain was grown on a minimal medium lacking leucine and tryptophan to select for both plasmids (DO-2) and on DO-2 without histidine to select for interaction (DO-3). As a positive control, this strain was observed to grow on the selective medium for dilutions ranging from 1 to 10^{-4} (Figure 19, lane 1). This result shows an interaction between HP1230 and HP1529 proteins, as previously identified using library screening (Rain *et al.*, 2001).

Two different plasmids were used for this study: (i) the pP6 vector which contain the GAL4 activation domain (AD) (Rain *et al.*, 2001). One of the HP1529 fragments (nucleotides 1-1374) obtained by screening the HP1230 protein was selected and used as prey in the pP6 vector fused to GAL4 AD; (ii) the p3H1 vector which contains the DNA-binding domain (DBD) of GAL4 and a methionine-regulated Met25 promoter (Tirode *et al.*, 1997, J. Biol. Chem. 272: 22995-22999). The HP1230 encoding sequence of 540 bp was sub-cloned from pB1-HP1230 into the *BamHI*/*PstI* sites of p3H1 as fusion protein with GAL4-DBD giving p3H1-HP1230. In addition, the WT SID1529 or SID1529* (N38D-V53L) or SID1529* (V53L) were sub-cloned from pP7-centro (NotI-BamHI) to the *NotI*/*BglII* sites of p3H1-HP1230 under the control of the Met25 promoter. Expression from the Met25 promoter is obtained in the absence of methionine. As negative control, we used a prey encoding the HP0875 protein. All PCR fragments and in frame fusions were checked by sequencing.

The pP6 and p3H1 derived-vectors were used to transform the Y187 and CG1945 yeast strains, respectively. Both strains were mated in YPD buffer (Yeast Peptone Dextrose; Bio 101, Inc) for 4 hours at 30°C and the resulting

diploid strain was selected on a minimal medium lacking leucine and tryptophane (DO-2). The interaction between proteins was observed in plates containing DO-2 deleted in histidine (DO-3) without methionine.

To assay whether The different versions of SID1529 can modulate this interaction, this protein was cloned in the p3H1-HP1230 vector under the control of the Met25 promoter. In the presence of the WT SID1529, growth was observed on DO3-Met medium thus showing that the WT SID1529 had little or no effect on the HP1230-HP1529 interaction (Figure 19, lane 2). In contrast, cells transformed with pP6-HP1529 and p3H1-HP1230-SID1529*(N38D-V53L) or p3H1-HP1230-SID1529* (V53L) were almost unable to grow on the selective medium (Figure 19, lanes 3-4). The growth of both strains in selective medium and in the presence of methionine (DO-3 + Met) was not affected, thus showing that the effect of SID1529* (N38D-V53L) or SID1529* (V53L) is a specific inhibition of the HP1230-HP1529 interaction (Figure 18, lanes 3-4). Taken together, these results clearly demonstrate that substitution of valine to leucine at position 53 of SID1529 led to complete inhibition of the HP1230-HP1529 interaction and confirm that SID1529 derivatived might have some potential as lead compounds to inhibit *Helicobacter pylori* growth.

Example 11 : Modulating compounds/PIM screening

The following results obtained from these Examples, as well as the teachings in the specification are set forth in the Tables below.

While the invention has been described in terms of the various preferred embodiments, the skilled artisan will appreciate that various modifications, substitutions, omissions and changes may be made without departing from the scope thereof. Accordingly, it is intended that the present invention be limited by the scope of the following claims, including equivalents thereof.

REFERENCES

Akerley, B.J., Rubin, E.J., Camilli, A., Lampe, D.J., Robertson, H.M., and Mekalanos, J.J. (1998) Systematic identification of essential genes by *in vitro* *mariner* mutagenesis. *Proc Natl Acad Sci USA* **95**: 8927-8932.

Alm, R.A., Ling, L.S., Moir, D.T., King, B.L., Brown, E.D., Doig, P.C., *et al.* (1999) Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori [published erratum appears in Nature 1999 Feb 25;397(6721):719]. *Nature* **397**: 176-180.

Bijlsma, J.J., Vandenbroucke-Grauls, C.M., Phadnis, S.H., and Kusters, J.G. (1999) Identification of virulence genes of Helicobacter pylori by random insertion mutagenesis. *Infect Immun* **67**: 2433-2440.

Blaser, M.J. (1993) *Helicobacter pylori*: microbiology of a 'slow' bacterial infection. *Trends Microbiol* **1**: 255-260.

Boyer, H.W., and Roulland-Dussoix, D. (1969) A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. *J Mol Biol* **41**: 459-472.

Cao, P., McClain, M.S., Forsyth, M.H., and Cover, T.L. (1998) Extracellular release of antigenic proteins by *Helicobacter pylori*. *Infect Immun* **66**: 2984-2986.

Chalker, A.F., Minehart, H.W., Hughes, N.J., Koretke, K.K., Lonetto, M.A., Brinkman, K.K., *et al.* (2001) Systematic identification of selective essential genes in *Helicobacter pylori* by genome prioritization and allelic replacement mutagenesis. *J Bacteriol* **183**: 1259-1268.

Chevalier, C., Thiberge, J.-M., Ferrero, R.L., and Labigne, A. (1999) Essential role of *Helicobacter pylori* g-glutamyltranspeptidase for the colonization of the gastric mucosa of mice. *Mol Microbiol* **31**: 1359-1372.

Colland, F., Rain, J.-C., Gounon, P., Labigne, A., Legrain, P., and De Reuse, H. (2001) Identification of the *Helicobacter pyllori* anti-s^{28} factor. *Molecular Microbiology* : in press.

Ferrero, R.L., Cussac, V., Courcoux, P., and Labigne, A. (1992) Construction of isogenic urease-negative mutants of *Helicobacter pylori* by allelic exchange. *J Bacteriol* **174**: 4212-4217.

GUY, B., C. HESSLER, S. FOURAGE, B. ROKBI, M.-J. QUENTIN MILLET (1999). Comparison between targeted and untargeted systemic immunizations with adjuvanted urease to cure *Helicobacter pylori* infection in mice. Vaccine **17** : 1130-1135.

Haas, R., Meyer, T.F., and van Putten, J.P. (1993) Aflagellated mutants of Helicobacter pylori generated by genetic transformation of naturally competent strains using transposon shuttle mutagenesis. *Mol Microbiol* **8**: 753-760.

Janvier, B., Grignon, B., Audibert, C., Pezennec, L., and Fauchère, J.L. (1999) Phenotypic changes of *Helicobacter pylori* components during an experimental infection in mice. **24**: 27-33.

Jenks, P.J., Chevalier, C., Ecobichon, C., and Labigne, A. (in press) Identification of non-essential *Helicobacter pylori* genes using random mutagenesis and loop amplification (RMLA). *Res Microbiol* .

Labigne, A. (1997) Random mutagenesis of the *H. pylori* genome. In *Helicobacter pylori protocols*. Mobley, C.L.C.a.H.M.T. (ed) Totowa, N. J.: Humana Press, pp. 153-163.

Labigne, A., Courcoux, P., and Tompkins, L. (1992) Cloning of Campylobacter jejuni genes required for leucine biosynthesis, and construction of leu-negative mutant of C. jejuni by shuttle transposon mutagenesis. *Research in Microbiology* **143**: 15-26.

Labigne-Roussel, A., Courcoux, P., and Tompkins, L. (1988) Gene disruption and replacement as a feasible approach for mutagenesis of *Campylobacter jejuni*. *J Bacteriol* **170**: 1704-1708.

McCarthy, M.A., Haebel, P.W., Torronen, A., Rybin, V., Baker, E.N., and Metcalf, P. (2000) Crystal structure of the protein disulfide bond isomerase, DsbC, from *Escherichia coli*. *Nat Struct Biol* **7**: 196-199.

Moskovitz, J., Rahman, M.A., Stassman, J., Yancey, S.O., Kushner, S.R., Brot, N., and Weissbach, H. (1995) *Escherichia coli* peptide methionine sulfoxide reductase gene: regulation of expression oand role in protecting against oxidative damage. *J Bacteriol* **177**: 502-507.

Odenbreit, S., Till, M., Hofreuter, D., Faller, G., and Haas, R. (1999) Genetic and functional characterization of the alpAB gene locus essential for the adhesion of Helicobacter pylori to human gastric tissue. *Mol Microbiol* **31**: 1537-1548.

Ogura, K., Maeda, S., Nakao, M., Watanabe, T., Tada, M., Kyutoku, T., *et al.* (2000) Virulence factors of *Helicobacter pylori* responsible for gastric diseases in monoglian gerbil. *J Exp Med* **192**: 1601-1610.

Parsonnet, J., Friedman, G.D., Vandersteed, D.P., Chang, Y., Vogelman, J.H., Orentreich, N., and Sibley, R.K. (1991) *Helicobacter pylori* infection and the risk of gastric carcinoma. *N Engl J* **325**: 1127-1131.

Parsonnet, J., Hansen, S., Rodriguez, L., Gelb, A.B., Warnke, R.A., Jellum, E., *et al.* (1994) *Helicobacter pylori* infection and gastric lymphoma. *N Engl J Med* **330**: 1267-1271.

Rain, J.-C., Selig, L., de Reuse, H., Battaglia, V., Reverdy, C., Simon, S., *et al.* (2001) The protein-protein interaction map of *Helicobacter pylori*. *Nature* **409**: 211-215.

Rashtchian, A. (1995) Novel methods for cloning and engineering genes using the polymerase chain reaction. *Current Opinion in Biotechnology* **6**: 30-36.

Salama, N., Guillemin, K., McDaniel, T.K., Sherlock, G., Tompkins, L., and Falkow, S. (2000) A whole-genome microarray reveals genetic diversity among *Helicobacter pylori* strains. *Proc Natl Acad Sci USA* **97**: 14668-14673.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular cloning: a laboratory manual*. Cold Spring Harbor, N. Y.: Cold Spring Harbor Laboratory Press.

Seifert, H.S., Chen, E.Y., So, M., and Heffron, F. (1986) Shuttle mutagenesis: a method of transposon mutagenesis for *Saccharomyces cerevisae*. *Proc Natl Acad Sci USA* **83**: 735-739.

Skouloubris, S., Thiberge, J.M., Labigne, A., and De Reuse, H. (1998) The Helicobacter pylori UreI protein is not involved in urease activity but is essential for bacterial survival in vivo. *Infection & Immunity* **66**: 4517-4521.

Tirode, F., Malaguti, C., Romero, F., Attar, R., Camonis, J., and Egly, J.M. (1997) A conditionally expressed thirt partner stabilizes or prevents the formation of a transcriptional activator in a three-hybrid system. *J Biol Chem* **272**: 22995-22999.

Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., *et al.* (1997) The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* **388**: 539-547.

Wizemann, T.M., Moskovitz, J., Pearce, B.J., Cundell, D., Arvidson, C.G., So, M., *et al.* (1996) Peptide methionine sulfoxide reductase contributes to the maintenance of adhesins in three major pathogens. *Proc Natl Acad Sci USA* **93**: 7985-7990.

Yoshimoto, T., Higashi, H., Kanatani, A., Lin, X.S., Nagai, H., Oyama, H., *et al.* (1991) Cloning and sequencing of the 7 alpha-hydroxysteroid dehydrogenase gene from *Escherichia coli* HB101 and characterization of the expressed enzyme. *J Bacteriol* **173**: 2173-2179.

Zapun, A., Missiakas, D., Raina, S., and Creighton, T.E. (1995) Structural and functional characterization of DsbC, a protein involved in disulfide bond formation in *Escherichia coli*. *Biochemistry* **34**: 5075-5089.

TABLE 1

Complexes of interacting proteins

1-Bait ORF	2-Prey ORF	1-Bait ORF	2-Prey ORF	1-Bait ORF	2-Prey ORF	1-Bait ORF	2-Prey ORF
HP0338	HP0132	HP1542	HP1430	HP0336	HP1479	HP0547	HP0802
HP0338	HP0337	HP1542	HP1542	HP0336	HP1541	HP0662	HP0662
HP0339	HP0006	HP0452	HP0155	HP0697	HP0012	HP0662	HP1168
HP0339	HP0116	HP0452	HP1409	HP0697	HP0048	HP0661	HP0432
HP0339	HP0178	HP0452	HP1523	HP0697	HP0558	HP0661	HP0605
HP0339	HP0252	HP0336	HP0056	HP0697	HP0599	HP0661	HP0609
HP0339	HP0267	HP0336	HP0066	HP0697	HP0696	HP0661	HP0876
HP0339	HP0298	HP0336	HP0100	HP0697	HP0864	HP0661	HP1245
HP0339	HP0486	HP0336	HP0112	HP0697	HP1037	HP1492	HP0172
HP0339	HP0544	HP0336	HP0116	HP0697	HP1038	HP1492	HP0351
HP0339	HP0578	HP0336	HP0145	HP0697	HP1299	HP1492	HP0464
HP0339	HP0613	HP0336	HP0154	HP0697	HP1576	HP1492	HP0593
HP0339	HP0626	HP0336	HP0192	HP1423	HP0697	HP1492	HP1041
HP0339	HP0705	HP0336	HP0230	HP1423	HP0775	HP1492	HP1286
HP0339	HP0736	HP0336	HP0281	HP1423	HP1024	HP1492	HP1533
HP0339	HP0922	HP0336	HP0306	HP1423	HP1198	HP0667	HP0620
HP0339	HP0961	HP0336	HP0364	HP1423	HP1222	HP0667	HP0770
HP0339	HP1084	HP0336	HP0405	HP1423	HP1421	HP0667	HP1430
HP0339	HP1092	HP0336	HP0470	HP1423	HP1529	HP0667	HP1571
HP0339	HP1156	HP0336	HP0476	HP0691	HP0692	HP1073	HP0552
HP0339	HP1207	HP0336	HP0540	HP0691	HP1362	HP1073	HP0973
HP0339	HP1241	HP0336	HP0592	HP0224	HP0224	HP1073	HP1309
HP0339	HP1333	HP0336	HP0605	HP0224	HP0919	HP1073	HP1409
HP0339	HP1373	HP0336	HP0621	HP0224	HP1374	HP1198	HP0088
HP0339	HP1513	HP0336	HP0626	HP0224	HP1409	HP1198	HP0268
HP1427	HP0025	HP0336	HP0660	HP0692	HP0197	HP1198	HP0293
HP1427	HP0116	HP0336	HP0665	HP0692	HP0584	HP1198	HP0452
HP1427	HP0229	HP0336	HP0680	HP0692	HP0887	HP1198	HP0705
HP1427	HP0289	HP0336	HP0701	HP0071	HP0288	HP1198	HP0775
HP1427	HP0317	HP0336	HP0717	HP0071	HP0378	HP1198	HP0965
HP1427	HP0373	HP0336	HP0760	HP0071	HP1165	HP1198	HP1032
HP1427	HP0810	HP0336	HP0774	HP0071	HP1450	HP1198	HP1114
HP1427	HP0896	HP0336	HP0775	HP0071	HP1503	HP1198	HP1124
HP1427	HP0913	HP0336	HP0779	HP0071	HP1503	HP1198	HP1198
HP1427	HP0944	HP0336	HP0786	HP0714	HP0116	HP1198	HP1274
HP1427	HP1120	HP0336	HP0794	HP0714	HP0289	HP1198	HP1378
HP1427	HP1157	HP0336	HP0795	HP0714	HP0527	HP1198	HP1411
HP1427	HP1157	HP0336	HP0850	HP0714	HP0528	HP1198	HP1541
HP1427	HP1157	HP0336	HP0858	HP0714	HP0552	HP1076	HP0033
HP1427	HP1243	HP0336	HP0903	HP0714	HP0714	HP1076	HP0275
HP1427	HP1342	HP0336	HP0919	HP0714	HP0958	HP1076	HP0304
HP0218	HP0527	HP0336	HP0929	HP0714	HP1138	HP1076	HP0319
HP1301	HP0030	HP0336	HP0996	HP0714	HP1300	HP1076	HP0449
HP1301	HP0705	HP0336	HP1026	HP0714	HP1446	HP1076	HP0753
HP1301	HP0887	HP0336	HP1037	HP0428	HP0284	HP1076	HP0841
HP1301	HP1409	HP0336	HP1132	HP0428	HP0705	HP1076	HP1544
HP1542	HP0289	HP0336	HP1236	HP0428	HP1350	HP1076	HP1563

HP1542	HP0289	HP0336	HP1241	HP0428	HP1409	HP1196	HP1409
HP1542	HP0289	HP0336	HP1266	HP1518	HP0116	HP1196	HP1514
HP1542	HP0601	HP0336	HP1274	HP1518	HP0705	HP1529	HP1230
1-Bait ORF	2-Prey ORF	1-Bait ORF	2-Prey ORF	1-Bait ORF	2-Prey ORF	1-Bait ORF	2-Prey ORF
HP1542	HP0609	HP0336	HP1325	HP1518	HP1065	HP0071	HP0278
HP1542	HP0610	HP0336	HP1332	HP1518	HP1489	HP0071	HP0570
HP1542	HP0610	HP0336	HP1434	HP0547	HP0299	HP0071	HP0775
HP1542	HP0740	HP0336	HP1452	HP0547	HP0601	HP0071	HP1340
HP1542	HP1273	HP0336	HP1452	HP0547	HP0687	HP0071	HP1409
HP0316	HP0621	HP0687	HP1503	HP0072	HP1489	HP0411	HP0304
HP0316	HP0895	HP0687	HP1506	HP0765	HP0764	HP0411	HP0484
HP0316	HP1350	HP0687	HP0174	HP0765	HP1024	HP0411	HP1019
HP1448	HP0107	HP0687	HP0687	HP0072	HP1489	HP0411	HP1326
HP0431	HP0547	HP0687	HP0759	HP0642	HP0642	HP0411	HP1453
HP0431	HP0705	HP0687	HP0942	HP1173	HP0528	HP0411	HP1470
HP0674	HP0033	HP0687	HP1185	HP1173	HP0595	HP0887	HP0610
HP0674	HP0862	HP0687	HP1212	HP1173	HP0610	HP0887	HP0887
HP0553	HP0553	HP0687	HP1510	HP1173	HP0816	HP0887	HP1157
HP0311	HP0312	HP0525	HP0389	HP1173	HP1553	HP0887	HP1464
HP1077	HP0304	HP0525	HP0513	HP1293	HP1198	HP1198	HP1032
HP1077	HP1472	HP0525	HP0525	HP1051	HP0435	HP1198	HP1218
HP0798	HP0798	HP0525	HP0550	HP1051	HP1513	HP1065	HP0918
HP0436	HP0010	HP0525	HP1409	HP1291	HP0278	HP1198	HP0264
HP0436	HP0072	HP0525	HP1411	HP1291	HP1283	HP1198	HP0371
HP0436	HP0858	HP0525	HP1451	HP1291	HP1409	HP1198	HP0615
HP0436	HP0887	HP0887	HP0459	HP1291	HP1553	HP1198	HP0655
HP0436	HP1019	HP0887	HP0610	HP0084	HP0392	HP1198	HP0793
HP0436	HP1513	HP0887	HP0699	HP0084	HP0950	HP1198	HP1002
HP0071	HP1409	HP0887	HP0887	HP0134	HP1012	HP1198	HP1048
HP0071	HP1550	HP0887	HP1157	HP0325	HP0246	HP1198	HP1069
HP1402	HP0420	HP0887	HP1460	HP0325	HP0278	HP0868	HP0088
HP1402	HP0621	HP0887	HP1464	HP0325	HP0338	HP0868	HP0327
HP1402	HP0657	HP0088	HP0011	HP0325	HP0399	HP0868	HP0869
HP1402	HP1355	HP0088	HP0088	HP0325	HP0449	HP0868	HP1142
HP0071	HP1409	HP0088	HP0221	HP0325	HP1321	HP0748	HP0048
HP0775	HP0033	HP0088	HP0437	HP0325	HP1409	HP0748	HP0100
HP0775	HP0056	HP0088	HP0464	HP0419	HP0321	HP0748	HP0695
HP0775	HP0231	HP0088	HP0792	HP0419	HP0419	HP0748	HP0736
HP0775	HP0285	HP0088	HP0974	HP0419	HP0862	HP0748	HP0748
HP0775	HP0378	HP0088	HP0988	HP0893	HP0862	HP0289	HP0289
HP0775	HP0621	HP0088	HP1045	HP0410	HP1080	HP0289	HP0289
HP0775	HP0645	HP0088	HP1096	HP0410	HP1198	HP0289	HP0610
HP0775	HP0705	HP0088	HP1198	HP1067	HP0392	HP0289	HP0887
HP0775	HP0775	HP0088	HP1535	HP0650	HP0012	HP0289	HP0922
HP0775	HP0792	HP1298	HP0012	HP0650	HP0272	HP0289	HP1038
HP0775	HP0965	HP1298	HP0278	HP0650	HP0317	HP0289	HP1542
HP0775	HP1063	HP1298	HP0449	HP0650	HP1026	HP0289	HP0289
HP0775	HP1274	HP1298	HP0522	HP0650	HP1245	HP0289	HP0289
HP0775	HP1355	HP1298	HP0528	HP0650	HP1547	HP0289	HP0610
HP0775	HP1409	HP1298	HP0696	HP0776	HP0067	HP0289	HP1355
HP1489	HP1489	HP1298	HP0705	HP0776	HP0278	HP0383	HP0547

HP1086	HP0019	HP1298	HP0770	HP0776	HP1378	HP0507	HP0172
HP1086	HP0278	HP1298	HP0887	HP0776	HP1409	HP0507	HP1409
HP1086	HP1409	HP1298	HP0950	HP0897	HP0422	HP0507	HP1489
HP0071	HP0189	HP1298	HP1379	HP0897	HP1362	HP1158	HP1158
HP0071	HP1072	HP1298	HP1403	HP0895	HP0316	HP1158	HP1329
1-Bait ORF	**2-Prey ORF**	**1-Bait ORF**	**2-Prey ORF**	**1-Bait ORF**	**2-Prey ORF**	**1-Bait ORF**	**2-Prey ORF**
HP0071	HP1165	HP0072	HP0017	HP0895	HP0426	HP1158	HP1409
HP0071	HP1486	HP0072	HP0073	HP0895	HP0894	HP1158	HP1527
HP0071	HP1503	HP0072	HP0156	HP0895	HP0895	HP0068	HP0070
HP0687	HP0705	HP0072	HP0216	HP0895	HP1409	HP0068	HP0118
HP0687	HP0889	HP0072	HP0437	HP0653	HP0247	HP0059	HP0067
HP0687	HP1072	HP0072	HP0610	HP0653	HP0653	HP0066	HP0066
HP0687	HP1329	HP0072	HP0680	HP0653	HP1012	HP0187	HP0020
HP0687	HP1487	HP0072	HP1138	HP0653	HP1409	HP0187	HP0231
HP0187	HP0254	HP1033	HP0379	HP1238	HP1488	HP0870	HP0748
HP0187	HP0705	HP0064	HP0063	HP1444	HP0010	HP0870	HP1157
HP0187	HP0862	HP1032	HP0643	HP1444	HP0088	HP0870	HP1243
HP0187	HP1414	HP1032	HP0818	HP1444	HP0099	HP0199	HP0278
HP0187	HP1513	HP1032	HP1122	HP1444	HP0392	HP0199	HP0356
HP0958	HP0353	HP1032	HP1198	HP1444	HP0760	HP0199	HP0449
HP0958	HP0522	HP1032	HP1316	HP1444	HP1116	HP0199	HP0891
HP0958	HP0582	HP0062	HP0062	HP1444	HP1332	HP0199	HP1198
HP0958	HP0714	HP0062	HP0103	HP1074	HP0347	HP0199	HP1313
HP0958	HP0884	HP0062	HP0170	HP1074	HP0593	HP0199	HP1314
HP0958	HP1428	HP0062	HP1156	HP1074	HP1278	HP0875	HP0874
HP0958	HP1462	HP0062	HP1409	HP1074	HP1316	HP0510	HP0347
HP0067	HP0069	HP0184	HP0123	HP1074	HP1399	HP0510	HP0510
HP0145	HP0116	HP0184	HP0238	HP1074	HP1493	HP1464	HP0132
HP0145	HP0281	HP0184	HP0354	HP1074	HP1497	HP1464	HP0740
HP0145	HP0468	HP0184	HP0609	HP0515	HP1409	HP1464	HP1198
HP0145	HP0887	HP0184	HP0655	HP0515	HP1537	HP1464	HP1464
HP0145	HP1274	HP0184	HP0705	HP0515	HP1561	HP1464	HP1510
HP0620	HP0620	HP0184	HP0793	HP1444	HP0088	HP0070	HP0068
HP0620	HP1409	HP0184	HP1409	HP1444	HP0392	HP0070	HP0070
HP0595	HP0025	HP0879	HP0033	HP1444	HP0422	HP0377	HP0033
HP0595	HP0060	HP0879	HP0066	HP1444	HP0655	HP0377	HP0145
HP0595	HP0595	HP0879	HP0252	HP1444	HP1332	HP0377	HP0183
HP0595	HP0870	HP0879	HP0401	HP1444	HP1493	HP0377	HP0194
HP0595	HP1052	HP0879	HP0407	HP1238	HP1238	HP0377	HP0224
HP0595	HP1430	HP0879	HP0421	HP1048	HP1198	HP0377	HP0285
HP0621	HP0609	HP0879	HP0626	HP0990	HP0085	HP0377	HP0399
HP0621	HP0610	HP0879	HP0649	HP0990	HP0116	HP0377	HP0486
HP0621	HP0862	HP0879	HP0703	HP0990	HP0321	HP0377	HP0854
HP1150	HP1451	HP0879	HP1248	HP0990	HP0663	HP0377	HP1409
HP0181	HP0807	HP0879	HP1274	HP0990	HP0727	HP0075	HP0075
HP1391	HP1041	HP0879	HP1379	HP0990	HP0909	HP0073	HP0067
HP0061	HP0066	HP0879	HP1422	HP0990	HP1213	HP0073	HP0232
HP0061	HP0978	HP0879	HP0012	HP0990	HP1327	HP0073	HP0705
HP0061	HP1409	HP0879	HP0116	HP0990	HP1333	HP0762	HP0272
HP1390	HP0118	HP0879	HP0178	HP0990	HP1345	HP0762	HP0506
HP1390	HP0120	HP0879	HP0275	HP1046	HP0162	HP0762	HP0655

HP1390	HP0154	HP0879	HP0281	HP1046	HP0182	HP0762	HP1409
HP1390	HP0325	HP0879	HP0321	HP1046	HP0220	HP0156	HP0156
HP1390	HP0351	HP0879	HP0483	HP1046	HP0509	HP0452	HP0109
HP1390	HP0528	HP0879	HP0578	HP1046	HP0583	HP0452	HP0155
HP1390	HP0751	HP0879	HP0639	HP1046	HP0746	HP0452	HP0278
HP1390	HP0809	HP0879	HP0666	HP1046	HP0750	HP0452	HP0705
HP1390	HP0816	HP0879	HP0699	HP1046	HP0760	HP0452	HP1282
1-Bait ORF	2-Prey ORF	1-Bait ORF	2-Prey ORF	1-Bait ORF	2-Prey ORF	1-Bait ORF	2-Prey ORF
HP1390	HP1012	HP0879	HP0735	HP1046	HP1024	HP0452	HP1409
HP1390	HP1081	HP0879	HP0821	HP1046	HP1197	HP1072	HP0096
HP1390	HP1188	HP0879	HP1012	HP1046	HP1312	HP1072	HP0792
HP1390	HP1379	HP0879	HP1160	HP1046	HP1464	HP1072	HP1321
HP1390	HP1486	HP0879	HP1299	HP1046	HP1488	HP1072	HP1409
HP1390	HP1513	HP0879	HP1374	HP1045	HP0325	HP1072	HP0050
HP0838	HP1409	HP0879	HP1459	HP1045	HP0750	HP1072	HP0231
HP1033	HP0599	HP0879	HP1513	HP1045	HP1205	HP1072	HP0263
HP1033	HP1513	HP0879	HP1567	HP0870	HP0317	HP1072	HP0705
HP1033	HP0651	HP1238	HP1238	HP0870	HP0488	HP1072	HP0836
HP1033	HP1105	HP1238	HP0655	HP0870	HP0595	HP1072	HP1430
HP0602	HP1402	HP1015	HP1521	HP0962	HP0617	HP1255	HP0687
HP0608	HP0175	HP0167	HP1035	HP0962	HP0621	HP1255	HP0969
HP0608	HP0589	HP1014	HP1014	HP0962	HP0655	HP1255	HP0969
HP0608	HP0626	HP1014	HP1409	HP0962	HP0750	HP1255	HP1072
HP0608	HP0919	HP0047	HP0047	HP0962	HP0996	HP1255	HP1077
HP0232	HP0605	HP0047	HP0048	HP0962	HP1019	HP1255	HP1114
HP0849	HP0056	HP0047	HP0695	HP0962	HP1024	HP1255	HP1290
HP0849	HP0066	HP1497	HP1074	HP0962	HP1025	HP1255	HP1550
HP0849	HP0109	HP1496	HP0769	HP0962	HP1188	HP0978	HP0979
HP0849	HP0116	HP0045	HP0045	HP0962	HP1274	HP0978	HP1583
HP0849	HP0132	HP0045	HP1409	HP0962	HP1353	HP0753	HP0030
HP0849	HP0197	HP0601	HP0547	HP0962	HP1379	HP0753	HP0033
HP0849	HP0213	HP0601	HP0753	HP0962	HP1409	HP0753	HP0115
HP0849	HP0354	HP0601	HP1154	HP0962	HP1478	HP0753	HP0197
HP0849	HP0370	HP0601	HP1377	HP0194	HP0592	HP0753	HP0492
HP0849	HP0373	HP0601	HP1409	HP0194	HP1377	HP0753	HP0747
HP0849	HP0417	HP1259	HP0006	HP0194	HP1488	HP0753	HP1035
HP0849	HP0452	HP1259	HP0025	HP0194	HP1489	HP0753	HP1098
HP0849	HP0466	HP1259	HP0025	HP0042	HP0017	HP0753	HP1274
HP0849	HP0476	HP1259	HP0026	HP0042	HP0051	HP0753	HP1454
HP0849	HP0512	HP1259	HP0033	HP0042	HP0054	HP0753	HP1485
HP0849	HP0539	HP1259	HP0212	HP0042	HP0278	HP0753	HP1523
HP0849	HP0569	HP1259	HP0231	HP0042	HP0381	HP0979	HP0977
HP0849	HP0570	HP1259	HP0252	HP0042	HP0419	HP0979	HP0979
HP0849	HP0583	HP1259	HP0280	HP0042	HP0723	HP0979	HP1278
HP0849	HP0593	HP1259	HP0317	HP0042	HP1133	HP0614	HP0614
HP0849	HP0618	HP1259	HP0354	HP0042	HP1318	HP0739	HP0016
HP0849	HP0621	HP1259	HP0405	HP0042	HP1392	HP0739	HP0379
HP0849	HP0717	HP1259	HP0585	HP1493	HP0033	HP0739	HP0453
HP0849	HP0760	HP1259	HP0615	HP1493	HP0043	HP0739	HP0527
HP0849	HP0971	HP1259	HP0643	HP1493	HP0051	HP0739	HP0547
HP0849	HP1037	HP1259	HP0705	HP1493	HP0116	HP0739	HP0621

HP0849	HP1041	HP1259	HP0779	HP1493	HP0281	HP0739	HP0855
HP0849	HP1048	HP1259	HP0792	HP1493	HP0435	HP0739	HP0887
HP0849	HP1069	HP1259	HP0830	HP1493	HP0483	HP0739	HP0810
HP0849	HP1084	HP1259	HP0862	HP1493	HP0501	HP0739	HP1056
HP0849	HP1129	HP1259	HP0965	HP1493	HP0669	HP0739	HP1157
HP0849	HP1132	HP1259	HP1012	HP1493	HP0841	HP0739	HP1279
HP0849	HP1213	HP1259	HP1232	HP1493	HP0977	HP0739	HP1357
HP0849	HP1299	HP1259	HP1259	HP1493	HP1002	HP1025	HP0668
HP0849	HP1332	HP1259	HP1299	HP1493	HP1025	HP1025	HP1025
1-Bait ORF	2-Prey ORF	1-Bait ORF	2-Prey ORF	1-Bait ORF	2-Prey ORF	1-Bait ORF	2-Prey ORF
HP0849	HP1374	HP1259	HP1342	HP1493	HP1138	HP1025	HP1114
HP0849	HP1379	HP1259	HP1411	HP1493	HP1198	HP1145	HP0182
HP0849	HP1393	HP1259	HP1422	HP1493	HP1350	HP1145	HP0289
HP0849	HP1400	HP1259	HP1430	HP1493	HP1411	HP1145	HP0400
HP0849	HP1452	HP0962	HP0030	HP1493	HP1533	HP1145	HP0550
HP0849	HP1489	HP0962	HP0088	HP0862	HP0607	HP1145	HP0621
HP0849	HP1493	HP0962	HP0111	HP0862	HP0621	HP1145	HP0669
HP0849	HP1553	HP0962	HP0244	HP0862	HP0791	HP1145	HP0690
HP1015	HP0299	HP0962	HP0285	HP0862	HP1474	HP1145	HP0705
HP1015	HP0547	HP0962	HP0410	HP0402	HP0825	HP1145	HP1010
HP1015	HP0607	HP0962	HP0426	HP0402	HP1130	HP0177	HP1409
HP1015	HP0687	HP0962	HP0490	HP0402	HP1234	HP0177	HP1508
HP1015	HP1435	HP0962	HP0528	HP0402	HP1478	HP0177	HP1550
HP1015	HP1490	HP0962	HP0537	HP1255	HP0607	HP0291	HP0279
HP0291	HP0291	HP1114	HP0919	HP1236	HP1148	HP0016	HP0791
HP0291	HP1409	HP0481	HP0281	HP1236	HP1263	HP0016	HP0969
HP0291	HP1541	HP0481	HP0298	HP0909	HP0587	HP0016	HP1017
HP0291	HP1550	HP0481	HP0416	HP0909	HP0607	HP0016	HP1183
HP0874	HP0875	HP0481	HP0421	HP0909	HP1081	HP0016	HP1267
HP1382	HP0025	HP0481	HP0494	HP0909	HP1521	HP0016	HP1450
HP1382	HP0182	HP0481	HP0645	HP0909	HP1561	HP0016	HP1490
HP1382	HP0269	HP0481	HP0736	HP1478	HP1478	HP1559	HP0010
HP1382	HP0317	HP0481	HP0741	HP1231	HP1247	HP1559	HP0056
HP1382	HP0705	HP0481	HP0779	HP0262	HP0355	HP1559	HP0087
HP1382	HP0896	HP0481	HP0971	HP0262	HP1318	HP1559	HP0116
HP1382	HP1142	HP0481	HP1000	HP1230	HP1230	HP1559	HP0116
HP1382	HP1182	HP0481	HP1025	HP1230	HP1529	HP1559	HP0116
HP1382	HP1237	HP0481	HP1266	HP0263	HP0056	HP1559	HP0120
HP1382	HP1243	HP0481	HP1406	HP0263	HP0486	HP1559	HP0132
HP1382	HP1342	HP0481	HP1435	HP0263	HP0605	HP1559	HP0145
HP0150	HP1037	HP0481	HP1513	HP0263	HP0779	HP1559	HP0355
HP0150	HP1067	HP1118	HP0822	HP0263	HP0792	HP1559	HP0364
HP0150	HP1198	HP0149	HP0103	HP0263	HP1238	HP1559	HP0379
HP0150	HP1371	HP0149	HP0196	HP0263	HP1325	HP1559	HP0406
HP0150	HP1409	HP0149	HP0331	HP0263	HP1332	HP1559	HP0466
HP0150	HP0449	HP0149	HP0527	HP0263	HP1489	HP1559	HP0488
HP0294	HP0100	HP0149	HP0547	HP0263	HP1513	HP1559	HP0589
HP0294	HP0289	HP0149	HP0693	HP0224	HP0250	HP1559	HP0606
HP0294	HP0377	HP0149	HP0775	HP0224	HP0920	HP1559	HP0615
HP0294	HP0399	HP0149	HP1034	HP0224	HP1076	HP1559	HP0621
HP0294	HP1318	HP0149	HP1079	HP0224	HP1198	HP1559	HP0660

HP0294	HP1409	HP0149	HP1115	HP0109	HP0101	HP1559	HP0663
HP0822	HP0651	HP0149	HP1248	HP0109	HP0488	HP1559	HP0669
HP0822	HP0725	HP0149	HP1332	HP0261	HP0281	HP1559	HP0680
HP0822	HP0822	HP0028	HP0214	HP0261	HP0537	HP1559	HP0710
HP0944	HP0033	HP0028	HP0354	HP0261	HP0583	HP1559	HP0780
HP0944	HP0224	HP0028	HP0527	HP0261	HP0605	HP1559	HP0795
HP0944	HP0695	HP0028	HP0547	HP0261	HP0632	HP1559	HP0817
HP0944	HP0944	HP0028	HP0607	HP0261	HP0655	HP1559	HP0858
HP0944	HP1318	HP0028	HP0791	HP0261	HP0750	HP1559	HP0929
HP0944	HP1529	HP0028	HP0969	HP0261	HP0887	HP1559	HP1026
HP0823	HP0100	HP0028	HP1041	HP0261	HP1019	HP1559	HP1041
1-Bait ORF	**2-Prey ORF**	**1-Bait ORF**	**2-Prey ORF**	**1-Bait ORF**	**2-Prey ORF**	**1-Bait ORF**	**2-Prey ORF**
HP0823	HP0255	HP0028	HP1072	HP0261	HP1124	HP1559	HP1075
HP0823	HP0352	HP0028	HP1222	HP0261	HP1274	HP1559	HP1116
HP0823	HP0488	HP0028	HP1450	HP0261	HP1350	HP1559	HP1143
HP0823	HP0919	HP0028	HP1470	HP0261	HP1489	HP1559	HP1187
HP0823	HP0969	HP0028	HP1503	HP0261	HP1513	HP1559	HP1207
HP0823	HP1024	HP0028	HP1550	HP0261	HP1517	HP1559	HP1236
HP0823	HP1116	HP0821	HP0705	HP0261	HP1543	HP1559	HP1266
HP0823	HP1409	HP0821	HP0821	HP0902	HP0588	HP1559	HP1274
HP0947	HP0096	HP0821	HP1114	HP0902	HP0887	HP1559	HP1325
HP0947	HP0116	HP0821	HP1409	HP0902	HP1409	HP1559	HP1393
HP0947	HP0589	HP0389	HP0696	HP0016	HP0016	HP1559	HP1434
HP0947	HP0621	HP0389	HP1262	HP0016	HP0038	HP1559	HP1452
HP0947	HP1012	HP0389	HP1400	HP0016	HP0228	HP1559	HP1453
HP0947	HP1205	HP0268	HP1198	HP0016	HP0439	HP1559	HP1470
HP0947	HP1373	HP1236	HP0026	HP0016	HP0687	HP1559	HP1472
HP1114	HP0651	HP1236	HP0347	HP0016	HP0697	HP1559	HP1479
HP1114	HP0705	HP1236	HP0355	HP0016	HP0759	HP1559	HP1489
HP1559	HP1493	HP1246	HP0326	HP1489	HP1489	HP0806	HP0134
HP0108	HP1399	HP1246	HP0407	HP0391	HP0099	HP0928	HP0928
HP0108	HP1541	HP1246	HP0886	HP0391	HP0392	HP0928	HP1409
HP0909	HP0996	HP1246	HP1035	HP1000	HP1409	HP1204	HP0514
HP0909	HP1041	HP1246	HP1244	HP0031	HP0006	HP1324	HP0010
HP0909	HP1081	HP1246	HP1460	HP0031	HP0066	HP0115	HP0317
HP0909	HP1241	HP0035	HP0035	HP0031	HP0281	HP0115	HP0753
HP0909	HP1256	HP1245	HP0650	HP0031	HP1041	HP0115	HP1157
HP0909	HP1321	HP1245	HP0717	HP0031	HP1513	HP1334	HP1329
HP1312	HP0083	HP1245	HP0816	HP0031	HP1567	HP1334	HP1409
HP1312	HP0282	HP1245	HP1024	HP0705	HP0254	HP0002	HP0002
HP1312	HP0413	HP1245	HP1523	HP0705	HP1406	HP0002	HP1200
HP1312	HP0705	HP0399	HP0399	HP0705	HP1541	HP0002	HP1409
HP1312	HP1046	HP0399	HP0621	HP0032	HP0032	HP0002	HP1451
HP1312	HP1316	HP0399	HP1048	HP0274	HP0066	HP0124	HP0377
HP1312	HP1409	HP0399	HP1430	HP0274	HP0518	HP0124	HP1409
HP1312	HP1430	HP1244	HP0857	HP0274	HP0951	HP0242	HP0242
HP1312	HP1523	HP1244	HP1246	HP0274	HP1294	HP0001	HP0050
HP0954	HP0045	HP0276	HP1263	HP1482	HP0102	HP0001	HP0355
HP0954	HP0466	HP0276	HP0012	HP1482	HP0259	HP0001	HP0792
HP0712	HP0152	HP0276	HP1489	HP1482	HP0466	HP0243	HP1409
HP0712	HP0197	HP0340	HP0011	HP1482	HP0489	HP1338	HP1298

HP0712	HP0278	HP0340	HP0096	HP1482	HP0645	HP1338	HP1338
HP0712	HP0281	HP0340	HP0231	HP1482	HP0760	HP1338	HP1409
HP0712	HP0662	HP0340	HP0340	HP1482	HP1115	HP1338	HP1523
HP0712	HP0714	HP0340	HP0695	HP1482	HP1369	HP0249	HP0012
HP0712	HP0736	HP0340	HP0730	HP0073	HP0073	HP0249	HP0099
HP0712	HP0836	HP0340	HP0821	HP0073	HP0232	HP0249	HP0174
HP0712	HP0850	HP0340	HP0834	HP0073	HP0259	HP0249	HP0214
HP0712	HP1160	HP0340	HP1088	HP1449	HP0615	HP0249	HP0415
HP0712	HP1470	HP0340	HP1138	HP1449	HP1019	HP0249	HP0527
HP0712	HP1489	HP0340	HP1198	HP1449	HP1332	HP0249	HP0547
HP0712	HP1513	HP0340	HP1274	HP1202	HP1068	HP0249	HP0610
HP0100	HP0123	HP0340	HP1362	HP1322	HP0013	HP0249	HP0787
HP0100	HP0132	HP0950	HP0030	HP1322	HP0649	HP0249	HP0922
1-Bait ORF	**2-Prey ORF**	**1-Bait ORF**	**2-Prey ORF**	**1-Bait ORF**	**2-Prey ORF**	**1-Bait ORF**	**2-Prey O**
HP0100	HP0278	HP0950	HP0116	HP1322	HP1488	HP0249	HP0969
HP0100	HP0585	HP0950	HP0192	HP0113	HP1041	HP0249	HP1041
HP0100	HP0617	HP0950	HP0304	HP1200	HP1037	HP0249	HP1154
HP0100	HP0707	HP0950	HP0394	HP1200	HP1199	HP0249	HP1157
HP0100	HP0710	HP0950	HP0884	HP1563	HP0785	HP0249	HP1377
HP0100	HP0726	HP0950	HP1025	HP1563	HP1080	HP0249	HP1409
HP0100	HP0830	HP0950	HP1198	HP0800	HP0433	HP0249	HP1487
HP0100	HP0852	HP0950	HP1438	HP0800	HP0687	HP0249	HP1503
HP0100	HP0898	HP1247	HP1231	HP0800	HP0800	HP0249	HP1543
HP0100	HP0919	HP1247	HP1353	HP0800	HP0801	HP0591	HP0582
HP0100	HP1050	HP0951	HP0116	HP0800	HP0924	HP0591	HP0610
HP0100	HP1157	HP0951	HP1409	HP0800	HP1267	HP0591	HP0663
HP0100	HP1379	HP0951	HP1411	HP0800	HP1460	HP0368	HP0409
HP0100	HP1411	HP1489	HP0001	HP0353	HP0353	HP0368	HP0435
HP0100	HP1458	HP1489	HP0351	HP0353	HP0406	HP0368	HP0857
HP0105	HP0105	HP1489	HP0540	HP0353	HP1420	HP0368	HP1293
HP0105	HP1409	HP1489	HP0705	HP0801	HP0152	HP0368	HP1356
HP0837	HP0705	HP1489	HP0797	HP0801	HP0800	HP0368	HP1410
HP0837	HP0836	HP1489	HP0841	HP0801	HP1513	HP0005	HP0088
HP1246	HP0121	HP1489	HP0862	HP0806	HP1198	HP0005	HP0116
HP0005	HP0175	HP1537	HP1189	HP1414	HP0522	HP0136	HP0983
HP0005	HP0278	HP1537	HP1245	HP1414	HP0655	HP0136	HP1119
HP0005	HP0317	HP1537	HP1325	HP1414	HP1124	HP0136	HP1195
HP0005	HP0583	HP1537	HP1430	HP1413	HP1024	HP0136	HP1205
HP0005	HP0615	HP1537	HP1489	HP1413	HP1409	HP0136	HP1241
HP0005	HP0640	HP1537	HP1541	HP1413	HP1413	HP0136	HP1393
HP0005	HP0657	HP1286	HP0231	HP1533	HP0010	HP0136	HP1430
HP0005	HP0665	HP1286	HP0625	HP1533	HP0033	HP0136	HP1510
HP0005	HP0714	HP1286	HP0833	HP1533	HP0351	HP0136	HP1553
HP0005	HP0977	HP1286	HP0934	HP1533	HP0381	HP0379	HP0706
HP0005	HP1274	HP1286	HP1119	HP1533	HP0402	HP0137	HP0010
HP0005	HP1350	HP1286	HP1222	HP1533	HP0944	HP0137	HP0066
HP0005	HP1409	HP1286	HP1379	HP1533	HP1430	HP0137	HP0109
HP0005	HP1411	HP1286	HP1492	HP1224	HP0278	HP0137	HP0197
HP0247	HP0117	HP1411	HP0033	HP1224	HP0781	HP0137	HP0244
HP0247	HP0247	HP1411	HP0351	HP1224	HP0976	HP0137	HP0247
HP0247	HP1316	HP1411	HP0397	HP1224	HP1409	HP0137	HP0289

619

		GATTTTACCGGTAATGGCGTTTTTGATA GCGTGAATTTCAACAAGGCTTATTACA AATTTCAAGGCACTGAAAATTCTTACAA TTTTAAAAACACGAACTTTTTAGCAGGG AATTTTAAGTTTCAGGGCAAGACCACCA TTGAAAAATCCGTTTTAAGCGACGCTTCT TACACTTTTGATGGCACGAATAACACCTT TACTGAAGACAAGTTTAATAATGGCTCG TTTAATTTTAGTCATGCAGAGCAGACAG ACGCTTTTAATAACAACTCGTTTAATGGC GGTTCGTTTAGTTTTAACGCCAAGCCAAGT AAATTTTAGTGGGAACTCGTTCAATGGGG GCGTGTTTAATTCAATAATACCCCTAAA GTCAGTTTCACTGATGACACTTTTAATGT GAATAACCAATTCAAAATAAATGGTACT CAAACAACTTTCACTTTCAATAAGGGCGT TGTTTTCAACATGCAAGGGCTTTTGAGCA GTTTAAGCGTAGGCACGACTTATCAATTG CTTA		

CLAIMS

What is claimed is:

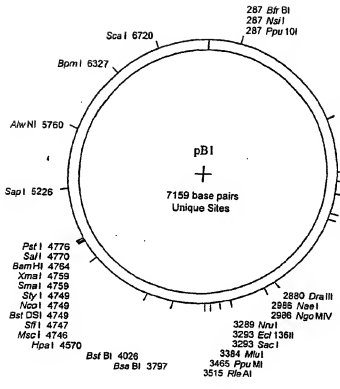
1. A complex of protein-protein interactions in *Helicobacter pylori* as defined in Table 1 and Table 8.
2. A complex of polynucleotides in *Helicobacter pylori* encoding for the polypeptides defined in Table 1 and Table 8.
3. A recombinant host cell expressing the interacting polypeptides in *Helicobacter pylori* defined in Table 1 and Table 8.
4. A method for selecting a modulating compound in *Helicobacter pylori* comprising:
   (a) cultivating a recombinant host cell on a selective medium containing a modulating compound and a reporter gene the expression of which is toxic for said recombinant host cell wherein said recombinant host cell is transformed with two vectors:
      (i) wherein said first vector comprises a polynucleotide encoding a first hybrid polypeptide and a DNA bonding domain;
      (ii) wherein said second vector comprises a polynucleotide encoding a second hybrid polypeptide and an activating domain that activates said toxic reporter gene when the first and second hybrid polypeptides interact;
   (b) selecting said modulating compound which inhibits the growth of said recombinant host cell.
5. A modulating compound obtained from the method of Claim 4.
6. A SID® polypeptide in *Helicobacter pylori* comprising the even SEQ ID Nos. 2 to 3256 in column 3 of Table 2, even SEQ ID Nos. 6590 to 6594 in Table 7 and even SEQ ID Nos. 6596 to 6644 in Table 8.
7. A SID® polynucleotide in *Helicobacter pylori* comprising the uneven SEQ ID Nos. 1 to 3255 in column 2 of Table 2, uneven SEQ ID Nos. 6589 to 6593 in Table 7 and uneven SEQ ID Nos. 6595 to 6643 in Table 8.
8. A vector comprising the SID® polynucleotide in *Helicobacter pylori* comprising the uneven SEQ ID Nos. 1 to 3255 in column 2 of Table 2,

uneven SEQ ID Nos. 6589 to 6593 in Table 7 and uneven SEQ ID Nos. 6595 to 6643 in Table 8.

9. A fragment of said SID® polypeptide according to Claim 6.
10. A variant of said SID® polypeptide according to Claim 6.
11. A fragment of said SID® polynucleotide according to Claim 7.
12. A variant of said SID® polynucleotide according to Claim 7.
13. A vector comprising the SID® polypeptide according to Claim 11.
14. A recombinant host cell containing the vectors according to Claim 8.
15. A pharmaceutical composition comprising a modulating compound of claim 5 and a pharmaceutically acceptable carrier.
16. A pharmaceutical composition comprising a SID® polypeptide of claim 6, and a pharmaceutically acceptable carrier.
17. A pharmaceutical composition comprising the recombinant host cells of claim 14 and a pharmaceutically acceptable carrier.
18. A protein chip comprising the polypeptides of claim 6.

pB1

Alias: pAS2DD
Application: 2HY (bait)
Backbone:
Specificity:
Selection: ampicillin
Constructed by:



Oligo 160

gagagtagtaacaaaggtc AAAGACAGTTGACTGTATCGCCG GAA TTT AT

Sfi I Sma I BamH I Sal I Pst I

G GCC ATG GAG GCC CCG GGG ATC CGT CGA CCT GCA GCC

Nco I

Oligo 161

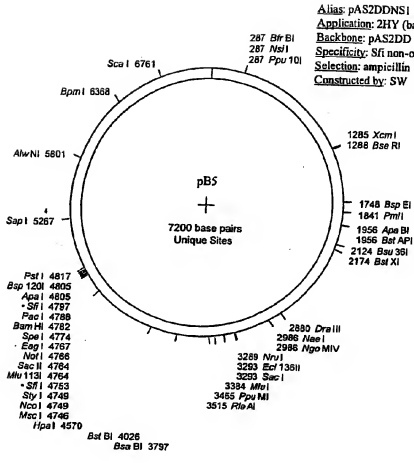
AAG CTA ATT ccgggcgaatttcttatg

Oligo 160 5' **GAGAGTAGTAACAAAGGTC** 3'
Oligo 161 5' **CATAAGAAATTCGCCCGG** 3'

FIGURE 1

2/20

pB5



Oligo 160

gagagtagtaacaaaggtcAAAGACAGTTGACTGTATCGCCG GAA TTT ATG

Sfi I **Sac II** **Spe I** **Bam HI**

GCC ATG GCC GCA GGG GCC GCG GCC GCA CTA GTG GGG ATC C

Nco I **Not I**

STOP **Sfi I** **Pst I**

TT AAT TAA GGG CCA CTG GGG CCC CTC GAC CTG CAG CCA

Pac I

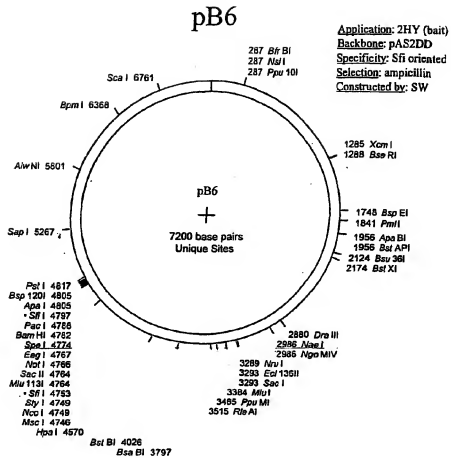
Oligo 161

AGC TAA TTccgggcgaatttcttatg

Oligo 160 5' **GAGAGTAGTAACAAAGGTC** 3'

Oligo 161 5' **CATAAGAAATTCGCCCGG** 3'

FIGURE 2


pB6
Application: 2HY (bait)
Backbone: pAS2DD
Specificity: Sfi oriented
Selection: ampicillin
Constructed by: SW
pB6
7200 base pairs
Unique Sites
287 Bfr BI
287 NsiI
287 Ppu 10I
Sca I 6761
Bpm I 6368
AlwNI 5801
Sap I 5267
Pst I 4817
Bsp 120I 4805
Apa I 4805
Sfi I 4797
Pac I 4786
BamHI 4782
Spe I 4774
Eag I 4767
Not I 4766
Sac II 4764
Mlu 113I 4764
Sfi I 4753
Sty I 4749
Nco I 4749
Msc I 4746
Hpa I 4570
Bst BI 4026
Bsa BI 3797
1285 Xcm I
1288 Bse RI
1748 Bsp EI
1841 PmlI
1956 Apa BI
1956 Bst API
2124 Bsu 36I
2174 Bst XI
2880 Dra III
2986 NaeI
2986 Ngo MIV
3289 NruI
3293 Eci 135II
3293 Sac I
3384 MluI
3485 Ppu MI
3515 Rle AI

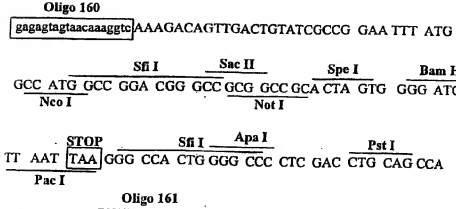
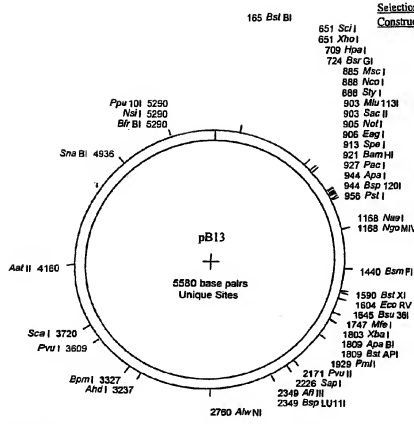

Oligo 160
gagagtagtaacaaaggtcAAAGACAGTTGACTGTATCGCCG GAA TTT ATG
Sfi I
Sac II
Spe I
Bam HI
GCC ATG GCC GGA CGG GCC GCG GCC GCA CTA GTG GGG ATC C
Nco I
Not I
STOP
Sfi I
Apa I
Pst I
TT AAT TAA GGG CCA CTG GGG CCC CTC GAC CTG CAG CCA
Pac I
Oligo 161
AGC TAA TTccgggcgaatttcttatg
Oligo 160 5' GAGAGTAGTAACAAAGGTC3'
Oligo 161 5' CATAAGAAATTCGCCCGG3'

FIGURE 3

4/20

pB13

Alias: pGBT9NS1
Application: 2HY (bait)
Backbone: pGBT9
Specificity: Sfi non-oriented
Selection: ampicillin
Constructed by: CR



Oligo 160

gagagtagtaacaaaggtc AAAGACAGTTGACTGTATCGCCG GAA TTT ATG

Sfi I Sac II Spe I Bam HI
GCC ATG GCC GCA GGG GCC GCG GCC GCA CTA GTG GGG ATC C
Nco I Not I

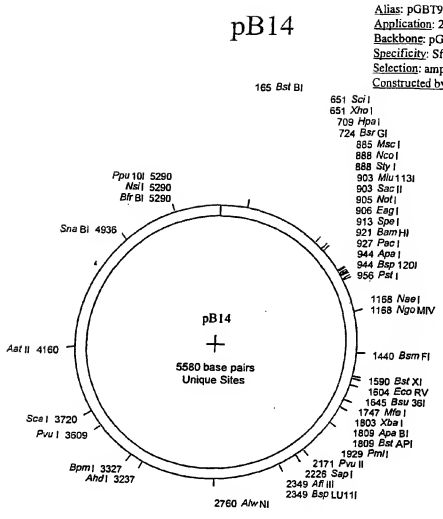
STOP Sfi I Pst I
TT AAT TAA GGG CCA CTG GGG CCC CTC GAC CTG CAG CCA
Pac I

Oligo 161

AGC TAA TT ccgggcgaatttcttatg

Oligo 160 5' **GAGAGTAGTAACAAAGGTC** 3'
Oligo 161 5' **CATAAGAAATTCGCCCGG** 3'

FIGURE 4



Oligo 160

gagagtagtaacaaaggtc AAAGACAGTTGACTGTATCGCCG GAA TTT ATG

Sfi I Sac II Spe I Bam HI

GCC ATG GCC GGA CGG GCC GCG GCC GCA CTA GTG GGG ATC C

Nco I Not I

STOP Sfi I Apa I Pst I

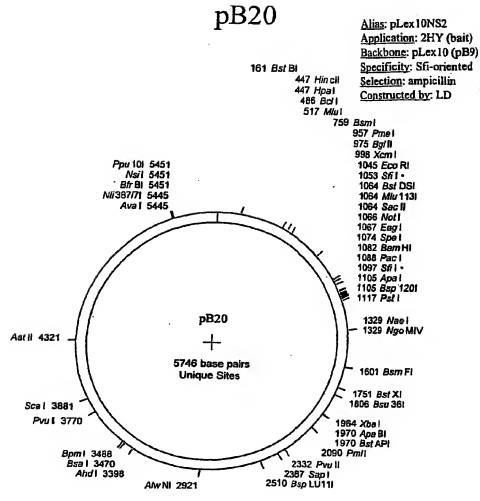
TT AAT TAA GGG CCA CTG GGG CCC CTC GAC CTG CAG CCA

Pac I

Oligo 161

AGC TAA TT ccgggcgaatttcttatg

FIGURE 5


pB20
Alias: pLex10NS2
Application: 2HY (bait)
Backbone: pLex10 (pB9)
Specificity: Sfi-oriented
Selection: ampicillin
Constructed by: LD
161 Bst BI
447 Hin cII
447 Hpa I
486 Bcl I
517 Mlu I
759 Bsm I
957 Pme I
975 Bgl II
998 Xcm I
1045 Eco RI
1053 Sfi I •
1064 Bst DSI
1064 Mlu 113I
1084 Sac II
1086 Not I
1087 Eag I
1074 Spe I
1082 Bam HI
1088 Pac I
1097 Sfi I •
1105 Apa I
1105 Bsp 120I
1117 Pst I
1329 Nae I
1329 Ngo MIV
1601 Bsm FI
1751 Bst XI
1806 Bsu 36I
1964 Xba I
1970 Ape BI
1970 Bst API
2090 Pml I
2332 Pvu II
2387 Sap I
2510 Bsp LU11I
Alw NI 2921
Bpm I 3488
Bsa I 3470
Ahd I 3398
Sca I 3881
Pvu I 3770
Aat II 4321
Ppu 10I 5451
Nsi I 5451
Blr BI 5451
Nli 3877I 5445
Ava I 5445
pB20
5746 base pairs
Unique Sites

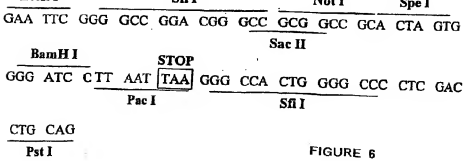
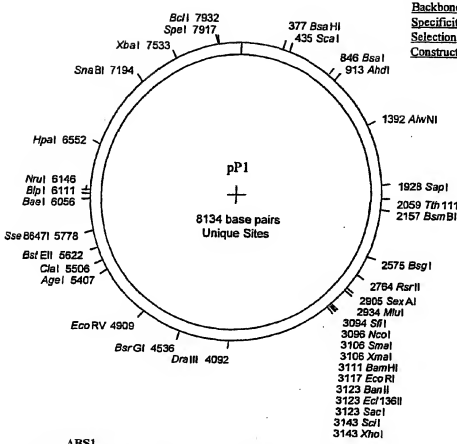

EcoR I
Sfi I
Not I
Spe I
GAA TTC GGG GCC GGA CGG GCC GCG GCC GCA CTA GTG
Sac II
BamH I
STOP
GGG ATC CTT AAT TAA GGG CCA CTG GGG CCC CTC GAC
Pac I
Sfi I
CTG CAG
Pst I

FIGURE 6

pP1

Alias : pACTIIst
Application: 2HY (prey)
Backbone: pACTII
Specificity:
Selection: ampicillin
Constructed by:



ABS1
cgtttggaatcactacagg GATGTTTAATACCACTACAATGGATGATGTATATAACTATCTATT

JC90
cgatgatgaagataccccaccaaa CCCAAAAAAAGAGATCTGTATGGCTTACCCATACGATGTTCCAG

Bgl II

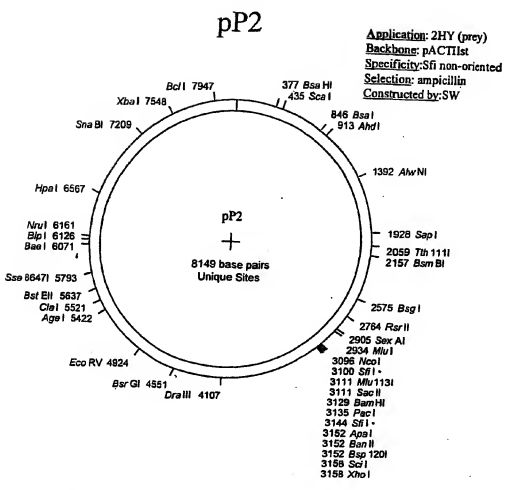
Sfi I Sma I BamH I
ATTACGCTAGCTTGGGTGGTCATATGGCC ATG GAG GCC CCG GGG ATC CGA ATT

Sac I Nco I Xho I Bgl II
CGA GCT CGA CTA GCT AGC TGA CTC GAG AGA TCT ATGAAT

cgtagatactgaaaaacccc GCAAGTT cacttcaactgtgcatcgtg caccatctcaatttc
162 ABS2 53

ABS1 5' CGTTTGGAATCACTACAGG 3'
JC90 5' CGATGATGAAGATACCCCACCAAA 3'
162 5' GGGGTTTTTCAGTATCTACG 3'
ABS2 5' CACGATGCACAGTTGAAGTG 3'
53 5' GAAATTGAGATGGTGCACGATGCAC 3'

FIGURE 7



ABS1

CG cgtttggaatcactacagg GATGTTTAATACCACTACAATGGATGATGTATATAACTATCTATT

JC90 Bgl II

cgatgatgaagataccccaccaaa CCCAAAAAAAGAGATCTGTATGGCTTACCCATACGATGTTCCAG

Sfi I Sac II

ATTACGCTAGCTTGGGTGGTCATATGGCC ATG GCC GCA GGG GCC GCG GCC GCA

Nco I

BamH I Pac I

CTA GTG GGG ATC CTT AAT TAA GGG CCA CTG GGG CCC CTC GAG AGA TCT

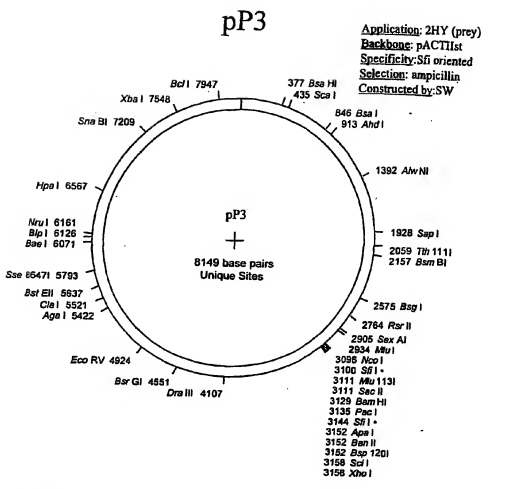
Stop

ATGAAT cgtagatactgaaaaaccc GCAAGTT cacttcaactgtgcatcgtg caccatctcaatttc

162 ABS2 53

ABS1 5' CGTTTGGAATCACTACAGG 3'
JC90 5' CGATGATGAAGATACCCCACCAAA 3'
162 5' GGGGTTTTTCAGTATCTACG 3'
ABS2 5' CACGATGCACAGTTGAAGTG 3'
53 5' GAAATTGAGATGGTGCACGATGCAC 3'

FIGURE 8



ABS1

CG cgtttggaatcactacagg GATGTTTAATACCACTACAATGGATGATGTATATAACTATCTATT

JC90

cgatgatgaagatacccccaccaaa CCCAAAAAAAGAGATCTGTATGGCTTACCCATACGATGTTCCAG

Bgl II

Sfi I Sac II

ATTACGCTAGCTTGGGTGGTCATATGGCC ATG GCC GGA CGG GCC GCG GCC GCA

BamH I Pac I Nco I

CTA GTG GGG ATC CTT AAT TAA GGG CCA CTG GGG CCC CTC GAG AGA TCT

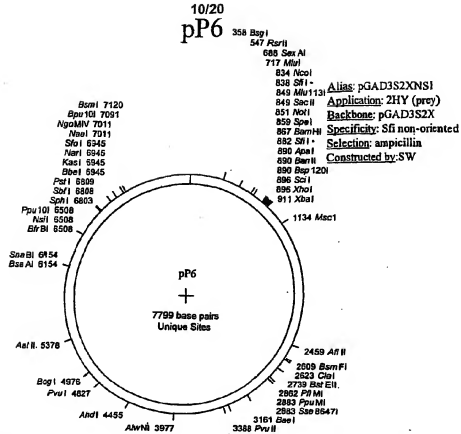
Stop

ATGAAT cgtagatactgaaaaacccc GCAAGTT cacttcaactgtgcatcgtg caccatctcaatttte

162 ABS2 53

ABS1 5' **CGTTTGGAATCACTACAGG 3'**
JC90 5' **CGATGATGAAGATACCCCACCAAA 3'**
162 5' **GGGGTTTTTCAGTATCTACG 3'**
ABS2 5' **CACGATGCACAGTTGAAGTG 3'**
53 5' **GAAATTGAGATGGTGCACGATGCAC 3'**

FIGURE 9



ABS1

cgtttggaatcactacagg GATGTTTAATACCACTACAATGGATGATGTATATAACTATCTATT

JC90

cgatgatgaagatacccaccaaa CCCAAAAAAAGAGATCCTAGAACTA

Sfi I Sac II Spe I Bam HI

GCC ATG GCC GCA GGG GCC GCG GCC GCA CTA GTG GGG ATC C

Nco I Not I

STOP Sfi I Xho I Xba I

TT AAT TAA GGG CCA CTG GGG CCC CTC GAG TAG CTA GTG TCT AGA

STOP STOP STOP

GGCCCGGTACCCAATTCGCCCTATAGTGAGTCGTATTACAATTCACTGGCCG TCGTTTTA

CAACGTCGTGACTGGGAAAACCCTGATCTATGAAT cgtagatactgaaaaacccc GCAA

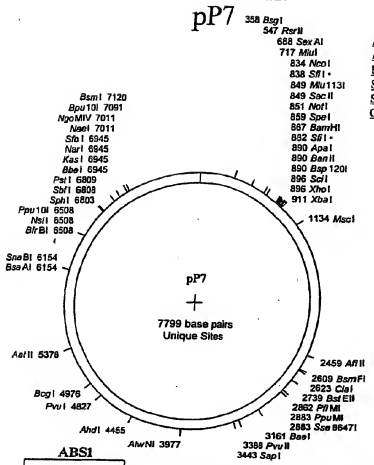
GTT cacttcaacgtgcatcgtg caccatctcaatttctttc

162

ABS2 53

ABS1 5' CGTTTGGAATCACTACAGG 3'
JC90 5' CGATGATGAAGATACCCCACCAAA 3'
162 5' GGGGTTTTTCAGTATCTACG 3'
ABS2 5' CACGATGCACAGTTGAAGTG 3'
53 5' GAAATTGAGATGGTGCACGATGCAC 3'

FIGURE 10



ABS1 5' CGTTTGGAATCACTACAGG 3'
JC90 5' CGATGATGAAGATACCCCACCAAA 3'
162 5' GGGGTTTTTCAGTATCTACG 3'
ABS2 5' CACGATGCACAGTTGAAGTG 3'
53 5' GAAATTGAGATGGTGCACGATGCAC 3'

FIGURE 11

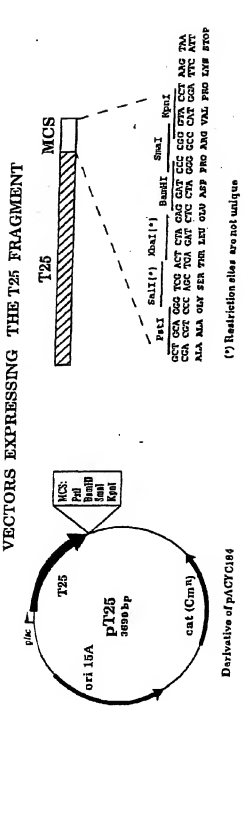

VECTORS EXPRESSING THE T25 FRAGMENT
T25
MCS
PstI SalI(*) XbaI(*) BamHI SmaI KpnI
GCT GCA GGG TCG ACT CTA GAG GAT CCC CGG GTA CCT AAG TTA
CGA CGT CCC AGC TGA GAT CTC CTA GGG GCC CAT GGA TTC AAT
ALA ALA GLY SER THR LEU GLU ASP PRO ARG VAL PRO LYS STOP
(*) Restriction sites are not unique
Derivative of pACYC184
pT25
3699 bp
ori 15A
cat (Cm^R)
ptac
MCS:
Pst
BamH
Sma1
Kpn1
T25
MCS
PstI SalI(*) XbaI BamHI SmaI KpnI
GCT GCA GGG TCG ACT CTA GAG GAT CCC CGG GTA CCT AAG TTA
CGA CGT CCC AGC TGA GAT CTC CTA GGG GCC CAT GGA TTC AAT
ALA ALA GLY SER THR LEU GLU ASP PRO ARG VAL PRO LYS STOP
(*) Restriction site is not unique
pKT25
3442 bp
ori 15A
kan (Km^R)
ptac
MCS:
Pst
BamH
Sac1
Kpn1
Derivative of pSU40

FIGURE 12

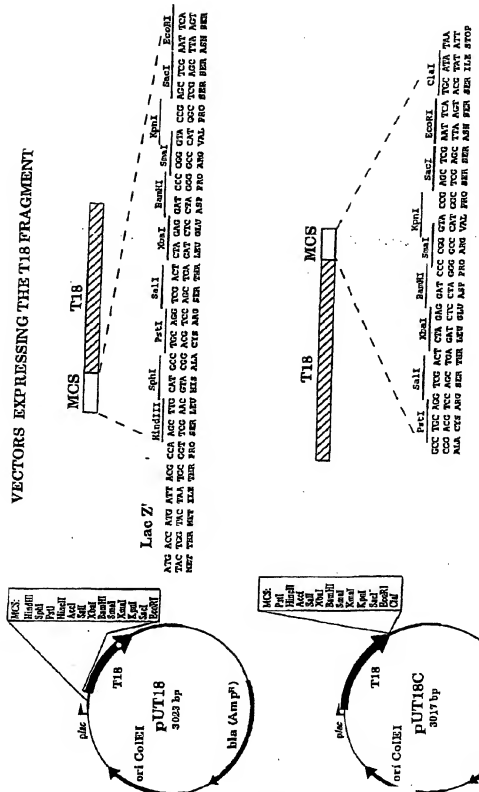

VECTORS EXPRESSING THE T18 FRAGMENT
MCS
T18
Lac Z'
pUT18
3023 bp
pUT18C
3019 bp
T18
MCS
plac
ori ColE1
bla (Amp^R)

FIGURE 13

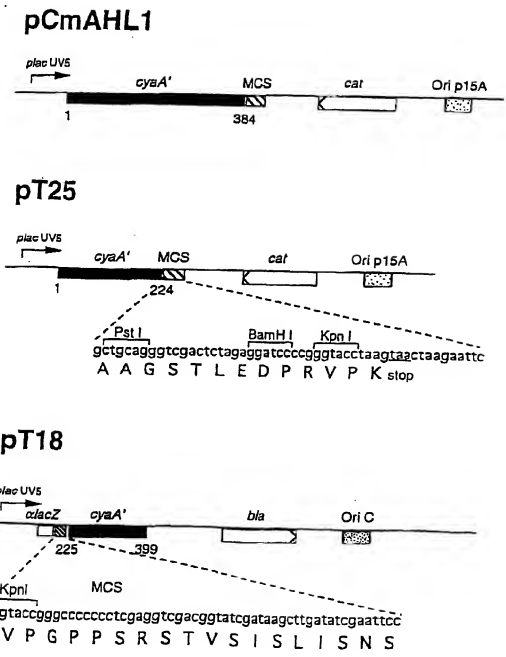
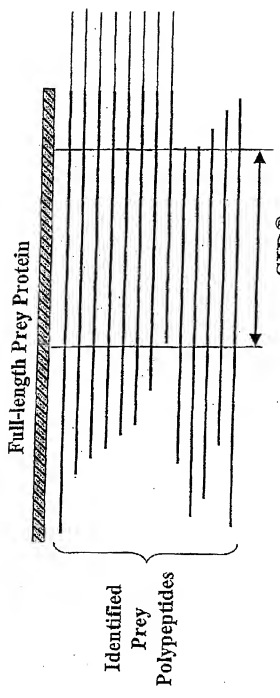
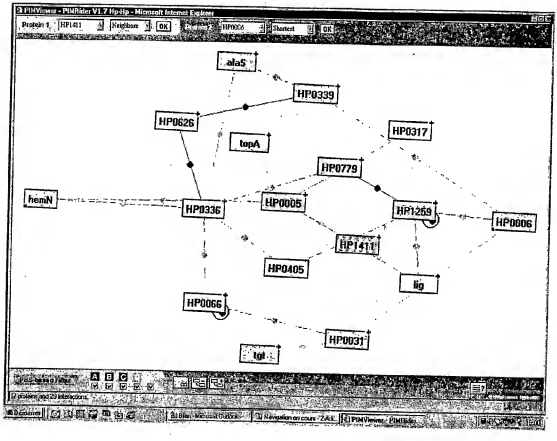

pCmAHL1
plac UV5
cyaA'
MCS
cat
Ori p15A
1
384
pT25
plac UV5
cyaA'
MCS
cat
Ori p15A
1
224
Pst I
BamH I
Kpn I
gctgcagggtcgactctagaggatccccgggtacctaagtaactaagaattc
A A G S T L E D P R V P K stop
pT18
plac UV5
cIacZ
cyaA'
bla
Ori C
225
399
KpnI
MCS
tgggtaccgggcccccctcgaggtcgacggtatcgataagcttgatatcgaattcc
W V P G P P S R S T V S I S L I S N S

FIGURE 14



Schematic representation of SID® determination

FIGURE 15



Exemple of Protein Interaction Map

FIGURE 16

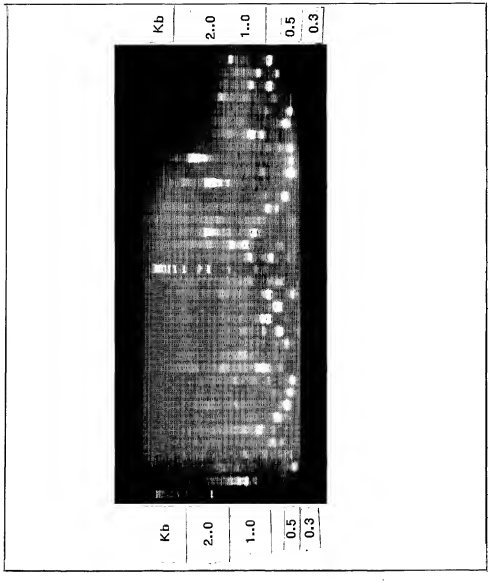


FIGURE 17

Procedure for the classification of the genes

H. pylori (HAS141) transformation:
4 independent patches: 1 µg super-coiled DNA/patch

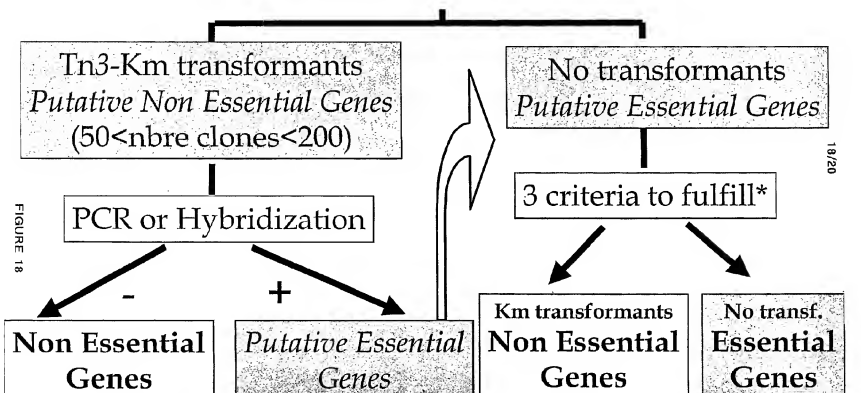


FIGURE 18

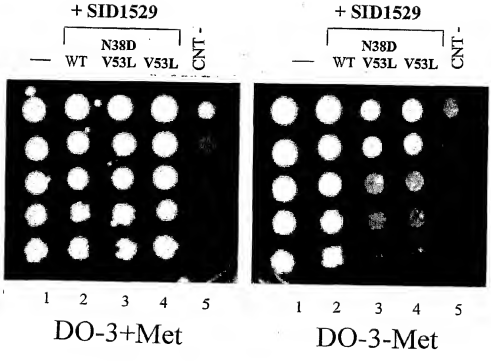
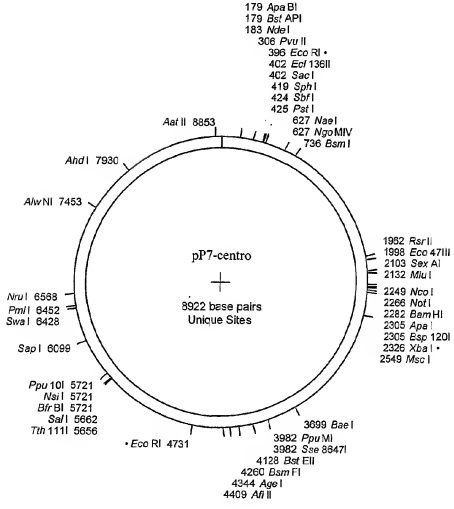

+ SID1529
N38D
— WT V53L V53L
CNT -
1 2 3 4 5
DO-3+Met
+ SID1529
N38D
— WT V53L V53L
CNT -
1 2 3 4 5
DO-3-Met

FIGURE 19

pP7-centro



ATGGCCGgAcGGGCC**GCGGCCGC**ACTAGTGG**GGATCC**TTAATTAAGGG

NotI BamHI

CCACTGGGGCCCCTCGAGTAGCTAGTGTCTAGAGGCCCGGTACCCAATT

FIGURE 20